Shobhana L. Chelliah
Willem J. De Reuse

Handbook of Descriptive Linguistic Fieldwork

Springer

Handbook of Descriptive Linguistic Fieldwork

Shobhana L. Chelliah • Willem J. de Reuse

Handbook of Descriptive Linguistic Fieldwork

Springer

Shobhana L. Chelliah
University of North Texas
Denton, TX 76203-1307
USA
chelliah@unt.edu

Willem J. de Reuse
University of North Texas
Denton, TX 76203-1307
USA
willemdereuse@my.unt.edu

ISBN 978-90-481-9025-6 e-ISBN 978-90-481-9026-3
DOI 10.1007/978-90-481-9026-3
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2010935299

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Acknowledgements

First, we owe a debt of gratitude to all the consultants and native speakers, too numerous to list, who of course ended up teaching us the most about the do's and don't's of fieldwork. Thank you all for your perceptive comments, telling smiles, looks, and frowns, and for your patience, when we asked you to repeat that one more time.

We are grateful to all the linguists and anthropologists, who, over the years, were our teachers, advisors and colleagues, and with whom we have had conversations or e-mails about fieldwork, or who made precious remarks, or were helpful in facilitating our fieldwork.

We want to thank in particular the linguists whom we interviewed formally for feedback on fieldwork: Emmon Bach, Balthazar Bickel, Anastasia Bonch-Osmolovskaja, Marc Brunelle, Bernard Comrie, Scott Delancey, Mark Donohue, Zarina Estrada Fernandez, Carol Genetti, Andrei Kibrik, Randy LaPolla, Jim Matisoff, David Peterson, Graham Thurgood, and George van Driem.

Colleagues or students originally from the University of North Texas with whom we talked about fieldwork include: Nancy Caplow, Patricia Cukor-Avila, Ivy Doak, Rosalia Dutra, Tim Montler, Kevin Mullin, Sadaf Munshi, Haj Ross, Peter Shillingsburg, and the late Walter Spitz.

Colleagues at the Research Centre for Linguistic Typology, at La Trobe University, or at Melbourne University, with whom we talked about fieldwork include: Alexandra Aikhenvald, Barry Blake, David Bradley, Rik de Busser, Seino van Breugel, Alec Coupe, Connie Dickinson, Gerrit Dimmendaal, Bob Dixon, Carola Emkow, Nicholas Evans, John Hajek, Friedel Frowein, Birgit Hellwig, Gerd Jendraschek, Rebecca Hanson, Randy LaPolla, Fiona Mc Laughlin, Elena Maslova, Stephen Morey, Rachel Nordlinger, Simon Overall, Regina Pustet, Tonya Stebbins, Nick Thieberger, Yvonne Treis, Ian Tupper, Pilar Valenzuela, Sheena van der Mark, Alejandra Vidal, and Roberto Zariquiey.

Our ideas about fieldwork were also enriched by discussions with faculty and students, from universities where we taught, and with teachers and fellow students at universities where we were students. Many of these people have now moved to other universities, but we mention them under the institution where we first met them. These institutions include: the Smithsonian Institution: Ives Goddard, Victor Golla, Sally McLendon, and William Sturtevant; the University of Kansas: Karen

Booker, Frances Ingemann, Karen Lupardus, John McLaughlin, Ken Miner, Robert Rankin, Kathleen Shea, and Akira Yamamoto; the University of Chicago: Ray Bucko, Lise Dobrin, Randy Graczyk, Eric Hamp, Paul Kroeber, Jerry Sadock, and Michael Silverstein; Ball State University: Elizabeth Riddle, and the University of Arizona: Jane Hill, Ken Hill, Marcia Inhorn, Muriel Saville-Troike, Rudy Troike, Mary Willie, and Ofelia Zepeda. We both studied with Tony Woodbury at the University of Texas at Austin. Long before it became fashionable, Tony understood the importance of fieldwork and grammar writing and promoted both at Austin. Without his encouragement and support we would never have experienced the field. We learned from other colleagues at UT: Brenda Boerger, Barbara McLeod, Frances Karttunen, Bob King, Carolyn Mackay, Joel Sherzer, Randy Valentine, and Jeff Williams.

Colleagues we talked to about fieldwork in South and Southeast Asia include Anvita Abbi, Gregory Anderson, Robbins Burling, Vineeta Chand, Gail Coelho, Scott DeLancey, Jerold A. Edmondson, Inga-Lill Hansson, Martine Mazaudon, Boyd Mikhailovsky, David Peterson, Martha Ratliff, Harimohon Thounaojam, Yashawanta Chungkam Singh, Martha Ratliff, Sohini Ray, and Nicolas Tournadre.

Colleagues or students with whom we talked about fieldwork in Siberia, Alaska, and Canada include: Andrea Berez, Knut Bergsland, Gary Donovan, Evgenij Golovko, Gary Holton, Steven Jacobson, Larry Kaplan, Jim Kari, Michael Krauss, Igor Krupnik, Molly Lee, Jeff Leer, Judy Maxwell, Bill Poser, Keren Rice, John Ritter, Siri Tuttle, and Anthony Woodbury.

Colleagues or students with whom we talked about fieldwork in the Southwest and Mexico include: Melissa Axelrod, Uldis Balodis, Keith Basso, Elizabeth Brandt, Joe Campbell, Karen Dakin, Emil Dobrescu, John Dunn, Philip Greenfeld, Frank Hardy, Martin Huld, Frances Karttunen, Paul Kroskrity, Jordan Lachler, Yolanda Lastra, Monica Macaulay, Pam Munro, Elinor Nevins, Natalie Operstein, Valentin Peralta, David Samuels, Irene Silentman, Muriel Saville-Troike, David Shaul, Siri Tuttle, Tom Willett, and Roberto Zavala.

Other colleagues or collaborators whose advice, guidance and words of wisdom we are grateful for include: Barry Alpher, Peter Austin, Claire Bowern, Bill Bright, Mike Cahill, Wallace Chafe, Ray DeMallie, Dan Everett, Aryeh Faltz, Anthony Grant, Marcia Haag, Mary Haas, Ken Hale, Leanne Hinton, Larry Hyman, Dell Hymes, Bruce Ingham, Paul Kroeger, Peter Ladefoged, Jack Martin, Ian Maddieson, Marianne Mithun, Paul Newman, Tom Payne, Kenneth Pike, Dennis Preston, David Rood, William Samarin, Jess Tauber, Allan Taylor, Peter Trudgill, Jan Ullrich, and Laurel Watkins.

We thank our parents, Raja Chelliah, Sita Chelliah, Joseph de Reuse, and Jozefa de Reuse-Claes, for their unfailing support.

Last but not least, we are grateful to Alec Coupe, who read the draft and provided valuable comments and corrections, to our hard-working language editor Nancy Caplow, and to the editors at Springer Publishers, who were a delight to work with and who were a model of patience and encouragement: Jolanda Voogd and Helen van der Stelt.

Contents

Chapter Synopsis of a Handbook of Descriptive Linguistic Fieldwork

Chapter 1: Introduction

Chapter 1 identifies the target audience for the Handbook. It reviews the current literature on fieldwork and closely related topics such as endangered language documentation. The authors report on how the Handbook is an original contribution to the field: it provides broad geographical coverage; a historical background of linguistic fieldwork; encyclopedic coverage of approaches and viewpoints on field sites, consultants, ethics, and methodology of data collection and analysis. The chapter concludes by laying out the organization of the book and provides a comprehensive reference list of other books on linguistic fieldwork.

Chapter 2: Definition and Goals of Descriptive Linguistic Fieldwork

Chapter 2 defines and discusses the primary, secondary, and ancillary goals of "descriptive linguistic fieldwork". The chapter quotes from several publications on fieldwork to reveal the diversity in opinions on what constitutes fieldwork, field sites, and consultants. The chapter ends with a comprehensive reference list on the topic.

Chapter 3: The History of Linguistic Fieldwork

Chapter 3 provides an in depth account of the history of linguistic fieldwork. It tracks Christian missionary fieldwork from early colonial times to the current day mission of the Summer Institute of Linguistics and Wycliffe Bible Translators; the writings of "gentlemen scholars" from the seventeenth to nineteenth centuries; commissioned word lists by wealthy or interested patrons; language descriptions based on emprisoned, enslaved, or hospitalized peoples. The chapter also discusses fieldwork based on collaboration with native consultants and fieldwork contemporaneous with traditions less supportive of fieldwork. The chapter ends with a discussion of the role of the Africanist fieldwork model and the current spate of fieldwork spurned by endangered language documentation and how these two influence our current understanding of what constitutes fieldwork. A comprehensive reference list is provided at the end of the chapter.

Chapter 4: Choosing a Language

Chapter 4 examines the factors which determine how a fieldworker selects a language to work on. Four main sources are identified and discussed: selection of the language is determined by an advisor or funding agency; the language community chooses the fieldworker; or the fieldworker selects the language. A comprehensive reference list is provided at the end of the chapter.

Chapter 5: Field Preparation: Research, Psychological, and Practical

In Chapter 5, three avenues of preparation for fieldwork are considered. First, there is the specific research preparation for the language and culture to be studied, which should be carried out in addition to general typological study (see Chapter 11). Second, is the psychological preparation by which the fieldworker learns from previous experience what to expect from the field situation and considers how his or her individual personality will react to and will deal with the pressures of the field. Third, the chapter deals with the practical aspects that must be taken care of before a fieldworker sets off to the field, including seeking funding; making contacts with a community of speakers; and purchasing and learning to use the right equipment. A comprehensive reference list is provided at the end of the chapter.

Chapter 6: Fieldwork Ethics: the Rights and Responsibilities of the Fieldworker

Chapter 6 is on the rights and responsibilities of the fieldworker to the host community, to the academic community, and to the self. Topics discussed include: the accurate and timely collection, description, and archiving of data; when called to do so by the community advocating for them and empowering and mobilizing them to be effective agents of language and cultural maintenance; exercising proper caution and providing appropriate guidance for students sent to the field; appropriately attributing data sources and acknowledging/honoring data ownership and appropriate safeguards against disallowed access to data; being aware of the consequences of fieldwork, many of which are unintended; having control over personal behavior in the field and with consultants; seeking the appropriate permissions from local authorities be they from the fieldworker's home institution, funding agency, or central, local, tribal governments at the field site. A comprehensive reference list is provided at the end of the chapter.

Chapter 7: Native Speakers and Field Workers

Chapter 7 deals with the selection of native speakers for a field project. Questions addressed include: how to begin looking for and hiring native speakers; establishing the role of speakers in the project depending on the speaker's individual characteristics (e.g. physical condition and age; gender; where they live; education and whether or not they are literate; personality traits; talent as a consultant; language proficiency; availability; and personal objectives in working on the field project); determining how many speakers to work with; exploring how good relations can be maintained with native speakers hired for the field project; and working with groups of speakers. In addition, the chapter looks at practical matters involved in dealing with consultants such as payment, gifts, and keeping track of native speakers through prompt cataloging of contacts. A comprehensive reference list is provided at the end of the chapter.

Chapter 8: Planning Session, Note Taking, and Data Management

Chapter 8 is a practical guide to organizing field sessions. The chapter advocates finding a convenient time and space for each session; building a flexible plan with clear objectives and a list of planned activities or tasks; preparation before the field session of necessary equipment and other materials needed to meet the stated objectives; and the recording of session data for further analysis by appropriate note taking, recording, data organization, and archiving practices. Included are suggestions for the internal organization of the field session from the introductory warm-up phase to the close. Interviewing techniques are reviewed with discussion of: how speakers may interpret an interview question; how fieldworkers should evaluate and react to native speaker responses; and how a native speaker's interest can be maintained during a field session. The chapter reviews ways of taking notes during each session, keeping track of data collected during the session (audio and video files, field notes, printed materials); and the archiving of these data. Suggestions are made about how fieldworkers can keep track of their finances. A comprehensive reference list is provided at the end of the chapter.

Chapter 9: Lexicography in Fieldwork

Chapter 9 distills from a wide variety of sources advice for the fieldworker on collecting words which can be used for wordlists and lexica or dictionaries. This chapter provides a methodology on how to create a wordlist by using a wordlist elicitation schedule of common words as well as extended culturally-specific wordlists and/or eliciting words from picture prompts, derived from texts; and dialect surveys. Advice is provided for organization of elicitation schedules by lexical category and semantic field. The chapter provides advice on methods of data cataloging and database management, along with notes on speaker input in finalizing bilingual dictionaries. A comprehensive reference list is provided at the end of the chapter.

Chapter 10: Phonetic and Phonological Fieldwork

Chapter 10 reviews the literature for advice on phonetic and phonological fieldwork. Topics covered include general preparation for such fieldwork such as learning how to transcribe and training the ear to perceive sounds not in the fieldworker's language. Guidance is also provided on how to organize word lists and short phrases for recording and; how to record such word lists for maximum success in phonetic and phonological analysis. It is also established that for this kind of fieldwork it is crucial to find the right kind and number of speakers to record, and to utilize native speaker input in determining sound distribution. We provide a detailed guide on how to elicit information on stress and tone. A comprehensive reference list is provided at the end of the chapter.

Chapter 11: Morphosyntactic Typology and Terminology

Chapter 11 will assist the beginning fieldworker in isolating the morphosyntactic, morphological, and syntactic characteristics of the target language. It provides a handy reference on topics such as: formal marking systems (head vs dependent, inverse, switch reference); lexical and grammatical categories; terminological issues in morphology such as the definition of the word; mechanisms of clause combination and transformation; and major hierarchies and scales relevant to syntax. An important feature of this chapter is the must-read lists of references provided with each of the described constructions.

Chapter 12: Grammar Gathering Techniques

Chapter 12 reviews approaches that fieldworkers can take when trying to understand the morphology and syntax of a language. The chapter is introduced by a discussion on directionality in grammar collection and grammatical theories that might influence data collection. After reviewing current terminology and classification of grammar gathering tasks and methods, details of data gathering techniques for morphosyntax, morphology, and syntax are discussed. The authors provide a detailed review of elicitation schedules along with a typology of such schedules. Also discussed is analysis controlled elicitation, that is, elicitation that is not based on a schedule such as: target language interrogation; stimulus-driven; target language manipulation; target language translation; target language construction and introspective judgment; reverse translation; review; ancillary; covert; and meta-elicitation. Specific techniques discussed for eliciting morphological facts include paradigm elicitation; bound and free morpheme elicitation; stem and root elicitation; and elicitation of noun, adjective and verb morphology. For gathering data on syntax, a discussion of the use of introspective syntactic judgments in grammatical description is provided in addition to elicitation by schedule, analysis controlled elicitation, and reverse translation. Each method is ranked for difficulty in terms of how hard the tasks are for the speaker to understand and perform and how difficult it is for the fieldworker to execute and interpret the results. A comprehensive reference list, including many online resources, is provided at the end of the chapter.

Chapter 13: Semantics, Pragmatics, and Text Collection

Chapter 13 reviews the place of semantics and pragmatics in descriptive linguistic fieldwork which are topics that are often given short shrift in field manuals. Included are a review of the basic definitions of word meaning (including synonomy, homophony, antonymy, and polysemy created through semantic change) sentence meaning; and pragmatics (including deixis, conversational implicature, presupposition, speech acts, and conversational structure). Finally this chapter provides a detailed discussion of text collection including a review of what is said in the literature about the advantages of text collection; types of texts that can be collected and how and where texts can be collected and analyzed (specifically creating translation; transcription; word-for-word translation; constituent analysis and free translations). A comprehensive reference list, including many online resources, is provided at the end of the chapter.

Chapter 1
Introduction

This book is a handbook, survey, and reference work for professional linguists and students of linguistics who intend to conduct descriptive linguistic fieldwork. Descriptive linguistic fieldwork (henceforth called linguistic fieldwork)[1] is the investigation of the structure of a language through the collection of primary language data gathered from interaction with native-speaking consultants.

As a handbook, it provides fieldworkers with detailed discussions of the theoretical, practical, and ethical issues involved in language selection, data collection, data management, interaction and work with consultants, and language analysis and description.

As a survey, the book covers past and present approaches and solutions to problems in the field, and the historical, political, and social variables associated with fieldwork in different areas of the world. The book also provides interested readers with access to topics through a detailed index and through comprehensive topical bibliographies at the end of each chapter.

In recent years there has been an increased interest in linguistic fieldwork. This is reflected in the publication, in just the last decade, of several book-length guides on the topic (Abbi 2001; Bowern 2008; Crowley 2007; Vaux and Cooper 1998; Vaux et al. 2007) and two book-length essay collections (Aikhenvald 2007; Newman and Ratliff 2001). Three more recent book-length collections of essays should be mentioned here, even though they are not considered guides to fieldwork: Gippert et al. (2006) which is on language documentation; and Ameka et al. (2006), and Payne and Weber (2006/2007) which are on grammar writing. These three collections refer perceptively to aspects of linguistic fieldwork, and the essays in them are almost all written by accomplished fieldworkers.[2]

[1] A detailed discussion of what is meant by descriptive linguistic field work as opposed to other sorts of linguistic fieldwork is provided in Chapter 2.

[2] For the sake of comprehensiveness, let us mention the three older book-length treatments of linguistic fieldwork: Samarin (1967), Bouquiaux and Thomas (1972), and Kibrik (1977). Samarin's *Field Linguistics* (1967) has served linguists well as a manual and is still a very comprehensive reference to the literature, but is now outdated in many technical, theoretical and sociolinguistic respects. Bouquiaux and Thomas's *Enquête et Description des langues à tradition orale* (1972) (and its more recent English translation Bouquiaux, Thomas, and Robert *Studying and Describing Unwritten Languages* (1992), include much helpful material such as sample questionnaires and a

S.L. Chelliah and W.J. de Reuse, *Handbook of Descriptive Linguistic Fieldwork*,
DOI 10.1007/978-90-481-9026-3_1,

Our contribution to this field is to provide a reference work that is broader in scope and coverage than existing volumes, from four points of view: geographical, historical, philosophical, and encyclopedic.

It is universally recognized that the experiences and the logistical, ethical, methodological, and analytical problems of descriptive fieldwork vary widely depending on the regions of the world where the fieldwork is conducted. This can be seen by comparing them: Abbi (2001) is geared to India, Bowern (2008) describes the Australian situation, and Crowley (2007) the Vanuatu (Melanesia) and Australian situations.[3] Vaux and Cooper (1998) does not cover geographical issues in any detail since it is a textbook designed for a 16-week semester field methods class with most of the examples taken from Armenian or Indian languages. Vaux et al. (2007), while not organized as a fieldwork class book, is an expansion of Vaux and Cooper (1998). *Linguistic Fieldwork*, edited by Newman and Ratliff (2001), and a special volume of *Sprachtypologie und Universalienforschung* on fieldwork edited by Aikhenvald (2007) have wider geographical coverage. However, because of the format of these works – they are collections of essays by a variety of field linguists – the relevant information on a single issue, specific field techniques for example, are not referenced in one place.

In this book, we attempt to provide geographical coverage that is as broad as possible. We draw on our personal and professional experience in India, Belgium, the American Southwest and Plains, Alaska, and the Canadian North. Additionally, based on personal interviews, we report on the experiences of seasoned fieldworkers from every continent. Finally, we direct readers to the published literature in order to provide resources relevant to their particular fieldwork situation.

This book pays more attention than its predecessors to the historical background of fieldwork. There are three things that the history of fieldwork can teach the fieldworker. First, and most obviously, the fieldworker can learn to avoid the mistakes of the past, which are more numerous, diverse, and imaginative than we care to believe. Second, and particularly in the case of endangered language documentation, the fieldworker often has to carry out a philological study of older documentation or fieldworker notes, and it is impossible to do philological work without a study of the historical context. Third, the study of the history of fieldwork helps in understanding why consultants, in many areas of the world, have the perceptions of linguistic fieldworkers that they have. These perceptions are typically due to past treatment (benign or otherwise) by colonial administrators; missionaries; aid, social, or medical workers; anthropologists; or other linguists. The areas of the world where the modern fieldworker is the very first outsider "in the field" are becoming very rare.

---

lengthy bibliography, but are also outdated as to content. Additionally, both of these works, especially Bouquiaux and Thomas, are geographically oriented towards the study of African languages. Kibrik's *The Methodology of Field Investigations in Linguistics* (1977) is outdated for the same reasons as Samarin, and is geographically oriented towards the languages of the former Soviet Union.

[3] As mentioned in the previous footnote, earlier books on fieldwork were geared towards Africa or the former Soviet Union.

Linguistic fieldwork is usually seen as conducted while living in a community far from the fieldworker's own and interacting with a typically rural or village culture very different from the fieldworker's own. We have a slightly different view about who counts as a consultant, and about the typical setting of linguistic fieldwork. We propose that there is such a thing as linguistic fieldwork in one's own community, or even with one's own relatives. For example, these days, community-based language preservation efforts are often guided by members of the community. More often than not, these documenters are non-speakers or semi-speakers of their heritage languages, and therefore in effect have to conduct fieldwork within their own communities or with their own relatives. Between these extreme contexts are the cases of linguistic fieldwork in urban offices, inner city apartments, or in Native reservation offices and schools, and so on. In our view, fieldwork issues in these varied environments are incorrectly considered to be only slightly different from the typical fieldwork situation. Also, we feel that fieldwork in the urban environment is mistakenly considered far less interesting than fieldwork in the typical village or rural context. Such "unexotic" fieldwork situations are more commonly encountered than discussed in the literature and should be investigated further.

Finally, the need for an encyclopedic handbook becomes obvious when one considers the recent surge of interest in endangered language preservation and documentation, as well as its connections with the more established issues of language politics and policies. This surge of interest has highlighted the importance of documentary and descriptive fieldwork in endangered languages, and numerous articles and websites on such fieldwork are now available. In this book we compile references to and distill the information from recent books, articles, and websites discussing linguistic fieldwork.

One recent encyclopedic work that is somewhat comparable in scope to this handbook is Dixon's two volume *Basic Linguistic Theory* (2010a, b). This work also views descriptive fieldwork, descriptive methodologies and analysis, and typology as stages on a path towards the goal of accurate descriptive linguistics. While there is very much we agree with in this book (and we will point the reader to it often), we do not always agree with the concept of Basic Linguistic Theory as formulated by Dixon, which we will discuss in some detail in Chapters 11 and 12.

This volume also includes discussion of material that is often omitted, or is covered in less detail. For example, we provide a full treatment of the investigation of grammar beyond the sentence: too often discussion of discourse and conversational analysis is omitted in accounts of fieldwork; however, these are crucial loci of grammatical information that cannot be found in non-continuous speech. We also discuss the collection, representation, management, and methods of extracting grammatical information from such data. We discuss the relationship between questionnaire-based elicitation, text-based elicitation, and philology, and the need for combinations of these methods.

We cannot claim to have covered every idea, saying, or opinion regarding linguistic fieldwork, but we are confident that we have gathered representative ideas and opinions all in one place.

Furthermore, because of its encyclopedic nature, this book is not only useful as a researchers' guide to take to the field, but also as a reference tool for beginners as well as for professional linguists to consult before, after, and in between field trips.

Many linguistics and anthropology departments offer field methods classes and encourage students to conduct original fieldwork to supplement more theoretical studies. This book can be useful to such students. Field methods classes are typically organized according to the complexity of the language being investigated and the level of the students in the class, and they tend to follow a predictable schedule, e.g. 4 weeks on phonology and 2 on syntax, etc. This book could constitute the sole class reading for a field methods class, with the division of readings being left up to the discretion of the instructor.

The remainder of the book is organized as follows. Chapter 2 defines descriptive linguistic fieldwork and states its goals. Chapter 3 discusses the history of linguistic fieldwork, illustrating how a historical understanding of fieldworkers, consultants, and linguistic descriptions is a necessary complement to the philological aspects of fieldwork. Chapter 4 presents insights on how fieldworkers select a region or language to work on. Chapter 5 is more practical in nature, and provides recommendations on how to prepare for fieldwork and what to take to the field. Chapter 6 addresses ethical issues. Chapter 7 focuses on how to set up working relationships with language consultants, and on what to expect in the field in terms of the personal and professional implications of fieldwork. Chapter 8 breaks down the structure of a typical fieldwork session, discussing how best to begin, manage, and end a fieldwork session. Special attention is given to the ordering of elicitation tasks, and to consultants' reception, attention and understanding of those tasks. Chapter 9 outlines methods for word list collection and discusses the special case of fieldwork leading to the creation of dictionaries. Chapter 10 is a guide to doing phonetic fieldwork, with extensive sections on the study of tone, stress and intonation, as well as step-by-step methodology on phonemic analysis. Chapter 11 summarizes the major terms and typological structures that a fieldworker should know about before going to the field. It also reviews relevant semantic, pragmatic and discourse-related concepts. Chapter 11 should be used along with Chapter 12, which provides a detailed overview of methods of elicitation and data gathering techniques. Here we suggest adoption of a method which consistently resets tasks according to the complexity of reasoning required by the consultant and the growing knowledge of the language on the part of the fieldworker. Finally, Chapter 13 focuses on semantic and pragmatic fieldwork, and provides a methodology for using texts in linguistic analysis. This is followed by a subject, language, and author index.

References

Abbi, Anvita. 2001. *A Manual of Linguistic Field Work and Indian Language Structures.* (Lincom Handbooks in Linguistics 17.) Munich: Lincom Europa.

Aikhenvald, Alexandra Y., ed. 2007. Focus on Linguistic Fieldwork. *Sprachtypologie und Universalienforschung* 60(1).

Ameka, Felix K., Alan Dench, and Nicholas Evans, eds. 2006. *Catching Language: The Standing Challenge of Grammar Writing.* (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Bouquiaux, Luc, and Jacqueline M. C. Thomas. 1972. *Enquête et Description des langues à tradition orale*. 1. L'enquête de terrain et l'analyse grammaticale; 2. Approche linguistique; 3. Approche thématique. Paris: Société d'Études Linguistiques et Anthropologiques de France.

Bouquiaux, Luc, Jacqueline M. C. Thomas, and James Robert. 1992. *Studying and Describing Unwritten Languages.* Dallas, TX: SIL International.

Bowern, Claire. 2008. *Linguistic Fieldwork Methodology: An Introduction to Data Elicitation Methods in the Field and the Classroom*. New York: Palgrave MacMillan.

Crowley, Terry. 2007. *Field Linguistics. A Beginner's Guide*. Oxford: Oxford University Press.

Dixon, R. M. W. 2010a. *Basic Linguistic Theory. Volume 1 Methodology*. Oxford: Oxford University Press.

Dixon, R. M. W. 2010b. *Basic Linguistic Theory. Volume 2 Grammatical Topics*. Oxford: Oxford University Press.

Gippert, Jost, Nikolaus P. Himmelmann, and Ulrike Mosel, eds. 2006. *Essentials of Language Documentation.* (Trends in Linguistics. Studies and Monographs 178.) Berlin/New York: Mouton de Gruyter.

Kibrik, A. E. 1977. *The Methodology of Field Investigations in Linguistics*. (Setting up the Problem.) (Janua Linguarum, Series Minor, 142.) The Hague: Mouton.

Newman, Paul, and Martha Ratliff, eds. 2001. *Linguistic Fieldwork*. Cambridge: Cambridge University Press.

Payne, Thomas E., and David J. Weber, eds. 2007. *Perspectives on Grammar Writing*. (Benjamins Current Topics 11.) Amsterdam/Philadelphia: John Benjamins. [Originally published as 2006, Special Issue of *Studies in Language* 30(2).]

Samarin, William J. 1967. *Field Linguistics. A Guide to Linguistic Field Work*. New York: Holt, Rinehart and Winston.

Vaux, Bert, and Justin Cooper. 1998. *Introduction to Linguistic Field Methods*. (Lincom Course books in Linguistics 1.) Munich: Lincom Europa.

Vaux, Bert, Justin Cooper, and Emily Tucker. 2007. *Linguistic Field Methods*. Eugene, OR: Wipf & Stock.

Chapter 2
Definition and Goals of Descriptive Linguistic Fieldwork

2.1 The Definition of Descriptive Linguistic Fieldwork

We define descriptive linguistic fieldwork as the investigation of the structure of a language through the collection of primary language data gathered through interaction with native-speaking consultants. Many other definitions emphasize the notion that the fieldworker must live like and with the native speakers of the language to be studied. For example, Everett (2001:168) defines linguistic fieldwork as:

> …the activity of a researcher systematically analyzing parts of a language other than one's native language (usually one the researcher did not speak prior to beginning fieldwork) within a community of speakers of that language, prototypically in their native land, living out their existence in the milieu and mental currency of their native culture.

A similar emphasis is also in Foley's discussion (2002:131):

> The ideal way to study the language of a traditional community is in situ, living with the village, learning as much of the social customs of the people as possible.

The same emphasis is present in Aikhenvald's (2007:5) definition as well:

> Linguistic fieldwork ideally involves observing the language as it is used, becoming a member of the community, and often being adopted into the kinship system.

Aikhenvald (2007:5–6) goes somewhat further than Everett and Foley, in that she distinguishes between "immersion fieldwork", which corresponds to her definition above, and "interview fieldwork", where the relationship between fieldworker and speaker is superficial and perhaps shorter, in that it is limited to interactions during fieldwork sessions. We hold that the success of the fieldwork endeavor is not based on whether fieldwork is of the "immersion" or "interview" style, but on whether it is intelligently or poorly conducted. In most fieldwork there is an "immersion" dimension, as the fieldworker tries to immerse her/himself in the community, as well as an "interview" dimension, when the fieldworker sits down with a consultant and asks questions. To be sure, no fieldworker has ever conducted fieldwork without asking questions. Equally true is the fact that "interview fieldwork" can be done with disastrous results, but then again, the same thing can be said of "immersion fieldwork", which can yield little analyzable data.

S.L. Chelliah and W.J. de Reuse, *Handbook of Descriptive Linguistic Fieldwork*,
DOI 10.1007/978-90-481-9026-3_2,

Everett, Foley and Aikhenvald are purists in this precise but romantic conception of fieldwork, much in the sense that the "participant observer" in the area of socio-cultural anthropology would consider himself or herself a purist in his/her field.

Other fieldworkers, such as Hyman (2001) and Samarin (1967:1–2), would consider the above definitions appropriate for prototypical fieldwork, but would agree that bringing the native speaker out of his/her milieu to another location, or working in an office is still considered fieldwork. While Crowley (2007:14–16) also holds that ideal fieldwork is in the community, he also accepts the possibility of fieldwork "at home".

Concerning the issue of prototypical versus less-prototypical fieldwork, Table 2.1 from Hyman (2001:21) provides a useful overview:

The prototype and the least fieldwork-like types described in this chart are sometimes caricatured by terms such as "dirty feet" linguistics (Crowley 2007:11–13) and "armchair" linguistics, respectively (Aikhenvald 2007:4, Crowley 2007: 11–13).

In this book, fieldwork is conceived of as having a slightly wider scope than what Everett, Foley, Aikhenvald, Samarin, Crowley, and Hyman have in mind. We define fieldwork both in terms of what it is and what it is not.

Descriptive linguistic fieldwork is:

1. Data collection for the purpose of the documentation and description of a language
2. Data collection through interaction with speakers
3. Data collection in situations where speakers are expected to use the language naturally

Descriptive linguistic fieldwork is *not*:

1. Data collection only through introspection
2. Data collection only through examination of written documents or written corpora
3. Data collection only through controlled lab experiments

Table 2.1 Prototypical versus less prototypical fieldwork (Reproduced from Table 1.1 in Hyman 2001)

	Fieldwork prototype	Fieldwork countertype	Least fieldwork-like
Elicitee	Other	Self	Introspection
Elicitor/observer	Self	Other	Secondary data
Distance	Far	Near	One's domicile
Setting	Small	Large	City, university
Duration	Long	Short	Brief stopover
Language	Exotic	Well-known	One's own
Subject matter	A language in its natural/cultural context	Language in general as a formal system	Abstract syntax
Data	Naturalistic	Controlled	Synthetic speech
Motivation	Languages-driven	Theory-driven	

We also argue that archiving, corpus-building and large lexicographic projects are not the concern of descriptive fieldwork. (See Section 9.3 for further comments on lexicography and fieldwork.)

Introspection, i.e. in some sense using oneself as a native-speaking consultant (discussed at length in Chapter 12), is not considered fieldwork in any discussion. However, in linguistic descriptions resulting from fieldwork, insights from fieldwork and from introspection are not always distinguished. Many descriptions by native-speaking linguists have been written using both introspection and speaker interaction; this interaction includes fieldwork with one's relatives, and fieldwork with others within their own communities. Some grammars of unwritten Flemish dialects were written this way by scholars who considered themselves dialectologists first and foremost. They were native speakers of the dialects they described, but nevertheless were superb descriptivist fieldworkers. Examples are Colinet (1896) on the phonetics and morphology of the Aalst dialect, Vanacker (1948) on the syntax of the Aalst dialect, and Pauwels (1958) on the Aarschot dialect. These descriptions, although quite conservative in that they are pre-phonemic, are nevertheless quite accurate and detailed.

There has been some debate on whether description based solely on the introspection of a native speaker can be considered fieldwork. For some, introspection is regarded as not only an efficient, but also the most reliable method for accessing a language's structure (See Chomsky 1957). The goal of the Chomskyan program is to build a model of linguistic competence. Since the structure of a language is present in each individual speaker, investigation into the competency of one fluent speaker should be a valid way to uncover the structure of that language, and a speaker could thus uncover his or her competency through introspection. There are some well-known examples of how a native speaker's introspective comments have been used for language description: see, for example, Sapir's (1933) work on the psychological reality of the phoneme, where a native speaker was encouraged to think about the distribution of sounds in his own language. In this way, fieldworkers often ask the native speaker to be introspective. See also Hale (1972) who has argued for the role of native speaker introspection in fieldwork.

There even exists a tradition within dialectology implying that introspection by speakers of an exotic or unwritten language counts as fieldwork. An example of this view is Basset (1951), who carried out fieldwork with Berber varieties in North Africa, and relied to some extent on introspection by natives.

There are other interactions with native speakers that we consider to be fieldwork. Sociolinguistic and dialectological pursuits – if involving interviews with native speakers – are considered fieldwork, following Lounsbury (1953:413–414) and Mosel (2001), and pace Munro (2003:130–131). Philological work – if carried out in consultation with native speakers – is also considered fieldwork. Several excellent descriptions have been written which combine fieldwork with research on earlier written sources, i.e. philology and epigraphy, as shown in Bowern (2008:4) and in Section 5.2 in this book.

Finally, we agree with Munro (2003:130–131) that the controlled lab experiments used by psycholinguists and language acquisition researchers are not fieldwork, but at the same time it needs to be acknowledged that controlled experimentation has a place, if a minor one, in fieldwork. Controlled experimentation has been particularly useful in phonetic fieldwork, as we will see in Chapter 10.

2.2 The Goals of Descriptive Linguistic Fieldwork

We consider that the goals of fieldwork depend on what sort of documents the fieldworker wants to produce. Not all fieldworkers state goals of fieldwork in terms of documents produced. For example, for Lounsbury (1953:414), fieldwork is a method "oriented toward a complete structural analysis of a language." For Vaux and Cooper (1999:17) the goal of fieldwork is to "elicit the maximum possible amount of reliable data in the minimum amount of time". Both goals are uniquely ambitious and uncomfortably vague. What indeed, is a "complete structural analysis?" What indeed, is the satisfactory "maximum amount of reliable data in the minimum amount of time"?

These are the sorts of questions we will attempt to answer in this book. In this chapter, we will also clarify what we mean by descriptive linguistic fieldwork. In the following sections we will distinguish three sorts of goals of linguistic fieldwork: primary goals (Section 2.2.1), secondary goals (Section 2.2.2), and ancillary goals (Section 2.2.3). The primary goals constitute what we will call descriptive linguistic fieldwork.

2.2.1 Primary Goals of Fieldwork

A European conception of descriptive linguistics distinguishes two methods of gathering data: (1) collecting a corpus of texts, which is part of what philologists traditionally do in their study of ancient written languages, and (2) interaction with a native speaker (Mosel 1987:10). Since for us fieldwork must involve interaction with a native speaker, only the second counts as real fieldwork.

In the American Boas–Sapir–Bloomfield tradition (Section 3.1), text collection and interaction with native speakers were not distinguished, since work was carried out on unwritten languages, and therefore all descriptive linguistics, including text gathering, originated in fieldwork, i.e. was based on interaction with native speakers. As a result, the European conception of descriptive linguistics as a cover term for two methods of data gathering can be discarded as too exclusive.

One can now distinguish (1) corpus collection of written documents, (2) corpus collection based on interaction with native speakers, (3) other activities based on interaction with native speakers. Activity (1) is part of the field of corpus linguistics, as well as of the field of philology. Activities (2) and (3) have given rise to the new field called "documentary linguistics", which can briefly be defined as the collection or gathering of linguistic data through a variety of methods and techniques, with a focus on reliability, representativity, and archivability. The field of "descriptive linguistics" is now conceived of as the analysis of language data gathered through activities (1) though (3). For some scholars, the goal of fieldwork should be documentation, whereas for other scholars the goal of fieldwork should not stop there, but should include descriptive linguistics as well. We will first discuss documentary linguistics as a goal, then descriptive linguistics as a goal, and then we will discuss the relationship between these two goals.

2.2.1.1 Documentary Linguistics

Documentation as a goal of fieldwork is, of course nothing new, since that was, after all, one of the goals of the Boas–Sapir–Bloomfield tradition (Woodbury 2003; Himmelmann 2006:14). At the time of this writing, documentary work is frequently being discussed because of the current attention to language endangerment issues (see Section 2.2.2.2).

Himmelmann (1998) is the foundational article arguing for a separation of documentary and descriptive fieldwork, within a broader field of descriptive linguistics (as originally defined in Section 2.2.1). We will argue in this chapter, and throughout this book, that a separation between documentary and descriptive fieldwork is not tenable, but first we will present in some detail the arguments for such a separation.

While Himmelmann (1998:163) recognizes that there is necessarily overlap in the area of the transcription of data in documentation and description, he argues that collection (i.e., documentary fieldwork) and analysis (i.e., descriptive fieldwork) are different activities in terms of result, procedure, and methodology. From a practical point of view, if collection and analysis are not distinguished, researchers will not pay sufficient attention to the activity of collecting. Secondly, when the documentary data are made available, they should be useful not only to people writing a descriptive grammar, but also to scholars in other disciplines such as anthropology, oral history, sociolinguistics, and discourse analysis. A grammatical description, on the other hand, is primarily useful only to grammarians and comparativists. Finally, description is different from documentation because there is no automatic procedure for deriving description from data, since depending on the underlying theoretical framework, different descriptions can and will result.

Lehmann (1999:1–2), holds a similar view of the distinction, and adds that since languages are dying faster than linguists can describe them, the only really urgent task is documentation. Lehmann distinguishes primary documentation, (i.e. a text corpus), from secondary documentation, (i.e. the description), and emphasizes that both must be accessible digitally. The documentation could be an "edited version of the field notes", and more ambitiously, what he calls a "radically expanded text collection", i.e. an annotated text collection, which should be a "record of the linguistic practices and traditions of a speech community" (Himmelmann 1998:165–166).

Further refinement of the definition of documentary linguistics is in Woodbury (2003). Woodbury's conception of documentary linguistics goes beyond a radically expanded text collection to include the full gamut of data obtained during fieldwork, from controlled or informal elicitations, commentary and grammaticality judgments by native speakers, to naturally-occurring speech recorded for its own sake. Woodbury (2007) further makes a convincing argument of the need for "thick translation", i.e. multiple levels and types of translations of one text.

Another account of what documentary linguistics is and what it should do is in Himmelmann (2006). This chapter recapitulates Himmelmann's (1998) views in a useful format, clarifies some terminology, and adds more historical context to the topic. It is, therefore, essential reading for the descriptive linguistic fieldworker. We do take exception to one idea in this important paper, which we quote here.

> It is a well-known fact that it is possible to base elaborate descriptive analyses exclusively on a corpus of texts (either texts written by native speakers or transcripts of communicative events) – and most good descriptive grammars are based to a large degree on a corpus of mostly narrative texts).
>
> (Himmelmann 2006:22)

We do not find this to be a well-known fact. While it is possible to produce a decent grammatical sketch of a language in this way, we argue in Chapter 12 and 13 that the dialogue between elicitation and texts is crucial to the writing of a good descriptive grammar.

On the whole, the above are convincing arguments for the existence of a separate field of documentary fieldwork. A question one can raise is whether field linguists can be collectors of corpora first and foremost. Traditionally, field linguists have not thought of themselves as collectors of corpora, even though they gather fieldnotes, texts and lexical material in a body that could be called a corpus. Most field linguists do not collect the sort of corpus that would be considered adequate for computational study of the sort done by corpus linguistics. Indeed, corpus linguistics, i.e. the analysis of previously collected corpora, is typically carried out with large world languages, such as English, French, or Hindi, with many speakers and extensive dialectal and stylistic variation, considerable written and recorded literature, and adequate funding and time devoted to their study. In the best pedagogical literature on these languages, there is a heavy reliance on data gathered from corpora. Corpus linguistics does not typically result from the activities of fieldworkers, since corpora typically consist of written data easily studied by computational methods, although they are increasingly transcripts from spoken data. Useful references on corpora are Johnson (2004), Meijs (1987), Oostdijk (1988), and Sampson (2002). Recent introductions to corpus linguistics include Kennedy (1998), McEnery and Wilson (1996), Teubert and Cermáková (2007), and Wynne (2005).

Documentary fieldwork is quite different, since interaction with speakers is assumed, there is always a certain urgency in gathering the data, and there is less concern over whether the data are statistically representative, properly sampled, and easily studied computationally. Documentary linguistics is a sort of emergency butterfly collecting, whereas corpus collecting would be a comprehensive butterfly collecting.[1]

There is no doubt that field linguists should increase efforts toward more representative corpus collecting when carrying out documentary fieldwork. Ultimately, when extensive corpora of all languages of the world have been gathered, the difference between corpus collecting adequate for corpus linguistics and documentary linguistics would become less important, but that goal is pie in the sky. We will probably never reach it.

Corpus collecting and documentary fieldwork are also different from the point of view of archiving. Archiving involves the procedures ensuring the preservation and continued availability of linguistic data. When collecting materials for a corpus, sampling techniques are important, and of course only what is sampled can be archived. One example of an archived linguistic corpus is the *Archivo de Lenguas Indígenas de México*, e.g. MacKay and Trechsel (2005) for Misantla Totonac.[2] When collecting materials in documentary fieldwork, the linguist is less selective, especially in the case of endangered languages where anything that can be collected is preserved archivally.[3] Examples of archives which contain the results of documentary fieldwork are the Archive of Indigenous Languages of Latin America (AILLA, University of Texas at Austin), the archive of the Alaska Native Language Center, (ANLC, University of Alaska, Fairbanks), the DOBES endangered languages archive (Max Planck Institute, Nijmegen, The Netherlands), and the Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC, Australia).[4]

We have pointed out that archiving implies preservation techniques. Lehmann (1999:10) points out that in other sciences such as archeology (artifacts) or zoology (preserved specimens), highly specialized techniques have been developed to preserve artifacts or specimens, and he laments the fact that such techniques do not yet exist in linguistics. He states: "We need to develop a culture of the linguistic datum and its processing." However, this point raises the question of whether a language can usefully be preserved like an archeological specimen, and the related ethical question of whether this is what native speakers or native speaker communities really want for their languages. Ethical questions relating to language description, documentation, archiving, and preservation are discussed in Chapter 6.

[1] As pointed out in Everett (2004), under the influence of Chomsky, field linguistics has disparagingly been compared to aimless "butterfly collecting". We urge field linguists to reclaim "butterfly collecting" as a positive term, and a particularly useful one if one wants to find out all about butterflies.

[2] The first 11 volumes of this archive, dealing with one Mexican indigenous language each, are now available on-line at http://www.colmex.mx/alim/.

[3] As the term "documentary" becomes more widespread in linguistics, so is the term "archival", used in new collocations such as: "archival phonetics" (Tuttle 2003), meaning using older sound recordings to carry out instrumental phonetics with them, and even "archival speakers" to designate the oldest, most conservative speakers of the Ainu language (DeChicchis 1995).

[4] All of these, and other archives less relevant to fieldwork, participate in the Open Language Archives Community (OLAC), (www.language-archives.org).

2.2.1.2 Descriptive Linguistics

According to the perceptive introduction to the edited volume on grammar writing by Evans and Dench (2006:3):

> The job of descriptive linguistics is to describe individual languages as perceptively and rigorously as possible, with maximal accountability to a naturalistic corpus of data ideally collected within a broad program of language documentation [...] to ensure that the full spectrum of language structures are represented.

We think that this definition also covers what descriptive fieldworkers should be doing, with the reservation, perhaps, that they should be doing this even if there is no "broad program of language documentation" in place yet. So, the goals of descriptive fieldwork are the writing of a comprehensive grammar, a collection of texts, and a dictionary, the so-called Boasian trilogy (Evans and Dench 2006:10–16). This trilogy was indeed an explicit goal of the Boas–Sapir–Bloomfield tradition, and is further discussed in Sections 3.1 and 9.1.

Lehmann's (1999:10) definition of description as a fieldwork goal is:

> Description of a language is an activity (and derivatively, its result), that formulates, in the most general way possible, the patterns underlying the linguistic data. Its purpose is to make the user of the description understand the way the language works.

According to Lehmann (1999:4–5), descriptions should aim at three things: (1) essential completeness, (2) intelligibility, and (3) adequacy.

"Essential completeness" does not mean that every detail is covered, but rather that all the main features of phonology, morphology, and syntax are covered, and that there is a dictionary and texts as well. Again, this was a goal explicitly stated by the Boas–Sapir–Bloomfield tradition. It fell by the wayside as post-Bloomfieldian structuralists tended to restrict themselves to phonology and morphology, and as their Chomskyan successors, in reaction, tended to restrict themselves to syntax.

"Intelligibility" implies that the description must be comprehensible to anyone with training in linguistics. Lehmann (1999:4–5) points out that tagmemic or transformational generative grammars written in the sixties are not good models, because they are no longer intelligible. In fact, the situation varies; the transformational account of Hidatsa (Siouan) syntax by Matthews (1965) is very hard to follow, but Lindenfeld's (1973) transformational syntax of Yaqui (Uto-Aztecan, northern Mexico) is still easy to read. The same argument can be made for some tagmemic accounts. Very readable tagmemic accounts, because they are commonsensical in presentation, are Bunn's (1974) grammar of Golin (Papua New Guinea), and De Wolf's (1997) grammar of Sonoran Mayo (Uto-Aztecan, northern Mexico).

Another matter of intelligibility is the avoidance of idiosyncratic terminology (Lehmann 1999:5, Mosel 2006:51). Idiosyncratic terminology became quite unwieldy in formal linguistics, particularly in later transformational-generative, minimalist, and optimality frameworks. In descriptivist milieus the situation is no better. For example, in the relatively small field of native North American language description, there are specialized terminologies for Algonquianists, Athabascanists, Eskimoanists, Iroquoianists, Muskogeanists, Salishists, Siouanists, and Uto-Aztecanists.

A well-established typological terminology is a strong desideratum, as further discussed in Section 10.5. A step towards terminology normalization has been taken by the E-MELD project's General Ontology for Linguistic Description (GOLD), available on-line at http://emeld.org/ontology-tree.cfm. It should still be a matter of discussion whether this terminology should be developed a priori, or a posteriori, i.e. departing from the specific usages of descriptivists.

"Adequacy" of course would include what Chomsky (1964) has called observational adequacy and descriptive adequacy, but for Lehmann (1999:5) it also means that the grammar should be written in such a general way as to be typologically comparable (Zaefferer 2006), but at the same time it should be specific enough "so that the uniqueness of the language is brought out".

2.2.1.3 On the Relationship Between Documentary and Descriptive Goals of Fieldwork

Informally, the relationship between documentary and descriptive goals (in terms of final products) can be set up as in Table 2.2.

Regarding the theoretical relationship between documentary and descriptive goals of fieldwork, there are three different points of view.

1. Himmelmann (1998, 2006) and Lehmann (1999, 2004) consider documentation and description to be theoretically independent, and consider that documentation should have priority as the goal of the fieldwork activity.
2. Woodbury (2003) also considers documentation and description to be theoretically independent, but considers documentation and description to have equal priority as the goal of the fieldwork activity.
3. Dixon (2007), republished in a slightly revised form in Dixon (2010:309–330), and Michael Krauss (p.c.) consider documentation and description to be theoretically dependent, and that description should have priority as the goal of the fieldwork activity. Dixon and Krauss disagree on the priorities within description, however. Dixon considers a reference grammar to be the priority, whereas Krauss considers a dictionary and text collections to be the priority.

Each of these points of view corresponds with different activities, and corresponds with different attitudes toward computerized data. Each of them have considerable merit, and the advantages and disadvantages of each will be briefly reviewed here.

Table 2.2 The relationship between documentary and descriptive goals

Type of data	Documentary	Descriptive
Word data	Word recordings	Dictionaries
Sentence data	Sentence recordings	Analyzed sentence examples
Discourse data	Text recordings	Analyzed texts
Integration of the above	–	Reference grammars

Himmelmann (1998) was the first proponent of a theoretical divide between the activities of documentation versus description, even though he admits that the dividing line is not always sharp in practice. Lehmann builds on this framework by further emphasizing the priority of documentation, as is clear from quotes such as:

> One should document a language in such a way that future linguists can derive a description from it.
>
> (Lehmann 1999:10)

> (…) let us call a sufficient documentation one on whose basis one can elaborate a description of the language. Now it is possible to come up with a sufficient documentation of a language within a few years. If the language then becomes extinct, it will still be possible to elaborate its description at leisure.
>
> (Lehmann 2004:63)

For Lehmann (2004:62, 63) the documentation contains the interface for the grammar, and the grammatical description is on a meta-level with respect to the documentation. In other words, fieldwork is primarily documentation, and description is a step beyond fieldwork. However, as reflected in our comments on Himmelmann's view (2006:22) quoted in Section 2.2.1.1, we do not believe that a comprehensive description can result from a study of documentary material without native speaker input.

The advantage of Lehmann's approach is that fieldworkers can concentrate on documentation, and can save the description for later. The disadvantage of this approach is that it is too optimistic in that it makes it seem like grammars and dictionaries can be computationally generated out of an annotated corpus. The processes would not be simple, but technological advances might make it possible to some extent. We have no way at present, however, to generate a comprehensive reference grammar out of a corpus. Good (2006a) has been studying reference grammars to determine to what extent they are similar to electronically generated (meta) databases. It is still too soon to know if investigations such as these will lead to computational grammar generation. In a paper about the ecology of documentary and descriptive linguistics (also worth reading for its candid assessment of relationships between computer programmers and descriptive linguists), Good (2006b) sees the ecology as a relationship between three individuals, the Archivist, the Collector and the User. If we assume that Good considers the Collector to be the Documentor, and the User to be the Describer or the heritage speaker, among others, then we have another view of the separation of description and documentation.

Woodbury (2003) shares Himmelmann and Lehmann's concern for the documentation of endangered languages, and a concern that documentation is under-theorized. Unlike Himmelmann and Lehmann, Woodbury does not view grammars as an endpoint of documentation, but rather as "part of the apparatus – the descriptive and explanatory material – that annotates the documentary corpus." Thus there is a dialectical relationship between the apparatus (or grammar) and the documentary corpus itself.

An influential voice for a distinction between documentation and description which has been instrumental in clarifying and expanding on Himmelman's and

Woodbury's points of view has been that of Peter K. Austin from the School of Oriental and African Studies, University of London. Austin is the editor of an impressive set of working papers entitled *Language Documentation and Description* (Austin 2003, 2004, 2005, 2007, 2008, 2009, 2010a). Austin has rightfully emphasized the complementarity of documentation and description in a series of survey articles (Austin and Grenoble 2007, Austin 2010b).

Against these points of view segregating description from documentary work, Dixon (2007) argues that it is neither possible nor advisable to consider documentary and descriptive fieldwork as distinct activities. Documenting is simply not enough, and the final product of fieldwork must be a reference grammar, a difficult and intellectually challenging task which can only be completed through the inductive generalizations of the fieldworker. Further support of this point of view is that when documentation and description are carried out in concert by the same linguist, the linguist gains a good overview of how the language works as a whole and both documentation and description benefit from this (Aikhenvald 2007 and Comrie 1988:5).

It is certainly significant that the two most recent accounts of grammar-writing, i.e. Ameka et al. (2006), and Payne and Weber (2007), largely contain contributions by fieldworkers, and that the recent manual of documentation, i.e., Gippert et al. (2006), also contains contributions by fieldworkers, and that the names of contributors to the descriptive and the documentation volumes broadly overlap. It is also significant that the collections of working papers mentioned earlier (Austin 2003, 2004, 2005, 2007, 2008, 2009, 2010a) also largely contain contributions by fieldworkers.

While we agree that documentation and description are theoretically distinct and complementary endeavors, our preference is with the approach that does not try to make too clear a segregation between the business of documentary linguistics and descriptive linguists. Keeping in mind the pressures of working against time to document a truly endangered language, we advocate fieldwork which leads to a comprehensive reference grammar and corpus of texts that can be used by linguists and speakers for a variety of purposes.

2.2.2 Secondary Goals of Descriptive Linguistic Fieldwork

Descriptive linguistic fieldwork also has secondary goals, which are instructional. One goal to impart native language Christian instruction (Section 2.2.2.1); another is to teach endangered languages to the next generation (Section 2.2.2.2). Neither of these goals follow from either documentary or descriptive goals. Both are to some extent controversial and involve a different set of researchers and team structure than do language documentation and description. Furthermore, we make no claim that both endeavors are equally valid from a humanist, moral, or ethical point of view; we just emphasize the fact that historically they have both been extremely important secondary goals.

2.2.2.1 Religious Instructional Goals

The goal of religious organizations such as the Summer Institute of Linguistics (SIL, nowadays called SIL International) and its missionary arm, the Wycliffe Bible Translators (WBT), is ultimately Bible translation. However, these organizations also encourage literacy among indigenous people who do not have a written language (Pittman 1948; Gudschinsky 1957). The reason for this is obviously that if the Bible is translated into an indigenous language, the indigenous people themselves have to be able to read it. Furthermore, literacy is conceived of as a valuable educational goal for the integration of indigenous peoples into the larger society. The relationship between literacy, literacy development, and fieldwork is somewhat controversial, since some indigenous communities might want to keep their language oral and are therefore opposed to literacy.

The issue of the need for Bible translation is much more controversial, of course, as discussed further in Section 3.2. In any event, SIL fieldwork has been praised by prominent non-SIL fieldworkers such as Comrie (1988) and Dixon and Aikhenvald (1999:2–3).

Table 2.3 below is a partial expansion of Table 2.1, showing the relationship between documentary, descriptive, and religious instructional goals. We hasten to point out that Table 2.3 is provided here for philological and historical purposes, since very few missionaries compile catechisms these days, and no one compiles confessionals[5] anymore.

The design of catechisms and confessionals was an important fieldwork activity carried out by missionaries in Spanish America. Examples of "confesionarios" are García (1760) for Coahuilteco of South Texas, discussed in Troike (1996:644–45), Beeler (1967) for Ventureño Chumash of California; and Ruz and Birrichaga (1997:289–299) for Zoque of Chiapas, Mexico. Examples of question and answer catechisms are Bausani (1974) for Chono of Chile; Beeler (1971:40–50) for a Yokuts variety of California; and Machoni (1877:215–221) for Lule of northern Argentina.[6]

Table 2.3 A comparison of documentary, descriptive and instructional religious goals

Type	Documentary	Descriptive	Instructional religious
Word data	Word recordings	Dictionaries	Dictionaries (including religious terminology)
Sentence data	Sentence recordings	Analysed sentence examples	Confessionals, and question-and-answer catechisms
Discourse data	Text recordings	Analysed texts	Doctrinal texts, Bibles
Integration of the above	–	–	Religious instructional texts in the target language

[5] Confessionals (Spanish "confesionarios") were bilingual phrasebook-like lists of set questions and answers, used by Spanish speaking Catholic missionaries in hearing confession from native converts.

[6] Except for Zoque, the languages mentioned in this paragraph are extinct.

2.2.2.2 Instructional Goals Relating to the Preservation of Endangered Languages

Since the seminal 1992 articles in *Language* (Craig 1992; England 1992; Hale 1992a, b; Jeanne 1992; Krauss 1992; Watahomigie and Yamamoto 1992); the literature on language endangerment has increased far more rapidly than has that on linguistic fieldwork. Edited book-length collections on the topic include Robins and Uhlenbeck (1991), Brenzinger (1998), Grenoble and Whaley (1998), Kasten (1998), Matsumura (1998), Ostler (1998), Fishman (2001), Sakiyama and Endo (2001), Bradley and Bradley (2002), Janse and Tol (2003), Sakiyama et al. (2004), Sakiyama (2004), De Dominicis (2006), Austin and Simpson (2007), Brenzinger (2007), Miyaoka et al. (2007), Moseley (2007), Harrison et al. (2008), and Austin and Sallabank (2010). Evans (2010) is a book for undergraduates, and is basically about endangered languages, but it is also particularly good at sharing the excitement of discoveries in the areas of language, culture, and thought; language and biology; language and the land, language and verbal art; and historical linguistics. Popular book-length accounts include Crystal (2000), Abley (2003), Dalby (2003), and Seay (2003). Other accounts, such as Nettle (1998), Nettle and Romaine (2000), and Harrison (2007) are somewhat elegiac about the ongoing language loss. Following this boom in literature on language endangerment, the literature on documentation aimed at preservation or stabilization (Cantoni 1996; Burnaby and Reyhner 2002), or teaching (Reyhner 1997) has also increased rapidly.

"Language preservation" or "language stabilization" include a variety of instructional activities aiming to prevent the break in the intergenerational transmission of a language, or to create a new generation of speakers in case the break in the intergenerational transmission has already occurred.[7] A useful overview of the terminological labels related to language preservation is Amery and Gale (2008:342). They prefer "language revival" as a cover term, and then distinguish three subtypes:

1. "Language revitalization" – the situation where there are maybe hundreds to a few older fluent speakers. This is a situation where the linguistic fieldworker can help with taking stock of the existing documentation, and can add to it.
2. "Language renewal" – the situation where there are no remaining speakers, but people remember some words and phrases. This is a situation where the linguistic fieldworker can help people jog their memories, for example by suggesting forms on the basis of what they know of related languages.
3. "Language reclamation" – the situation where nothing of the language is remembered, and the materials for relearning the language have to be based on historical documents. This is a situation where fieldworkers can be of no direct help. If the

[7] We focus here on the instructional activities included in "language preservation" or "language stabilization", because that is where the fieldworker can be most helpful. The fieldworker should always remain aware of the fact that "language preservation" or "language stabilization" also include activities such as language planning and language policy, and therefore that any "language preservation" or "language stabilization" effort has political causes and consequences.

fieldworker is good at philologically interpreting other people's fieldnotes (see Section 5.2), s/he can help indirectly in this way. However, for a descriptive fieldworker, work in situation (1) should always remain the highest priority, and work in situation (3) the lowest.

Practical advice related to language revival fieldwork is contained in the survey by Hinton and Hale (2001), in Hinton et al.'s (2002) manual, in Grenoble and Whaley's (2006) survey, and in Austin and Sallabank (2010). These works deal with documenting and describing a language with the ultimate goal of learning or relearning it. This literature also contains discussion of technical and orthographic issues related to language instruction. The best overview of the problems arising when doing fieldwork with speakers of endangered languages with the goal of writing instructional materials is Grinevald (2007). A good overview of multimedia teaching techniques for endangered languages, as derivable from fieldwork-based documentation, is in Nathan (2006), and an overview of orthography development is in Seifart (2006).

Table 2.4, also derived from Table 2.2, compares documentary, descriptive, and language instructional goals.

While not nearly as controversial as the religious goals, there have also been skeptical voices on the validity of these as goals for linguistic fieldwork (Ladefoged 1992; Newman 1998; Mufwene 1998). It is probably no coincidence that these voices are from Africanists. They were the first, as discussed in Section 3.6, to reflect critically on the goals of linguistic fieldwork, and have been among the first to voice skepticism about the current optimism in language endangerment related fieldwork. There is also a question of priorities: We are in agreement with Comrie (2007), who argues that documentary work on endangered languages should remain a higher priority than the revitalization of extinct or non-traditional varieties.

2.2.3 Ancillary Goals of Descriptive Fieldwork

In this section we discuss other types of linguistic fieldwork, which are not primarily descriptive. We consider descriptive fieldwork, in addition to its important goals which are valid in their own right, can also be ancillary to those other types of

Table 2.4 The relationship between documentary, descriptive, and language instructional goals

Type	Documentary	Descriptive	Language instructional
Word data	Word recordings	Dictionaries	Learner's dictionaries
Sentence data	Sentence recordings	Analyzed sentence examples	Phrasebooks
Discourse data	Text recordings	Analyzed texts	Primers or readers
Integration of the above	–	Reference grammars	Pedagogical grammars, textbooks, or multimedia learning methods

linguistic fieldwork. The use of the term "ancillary" is not intended to imply that the sorts of linguistic fieldwork described here are more or less important than descriptive fieldwork. It is just that some linguistic fieldwork is not descriptive, and that while the goals of such fieldwork are different, descriptive fieldwork practices will always be useful to help reach these goals.

2.2.3.1 Non-comparative Theoretical Goals

The goals of non-descriptive fieldwork can be to substantiate theoretical claims[8] regarding such concepts as Universal Grammar (Abbi 2001; Evans and Levinson 2009), the biologically hardwired language acquisition device, or the independence or relationship between form and function (Evans and Dench 2006:7–10). As Mosel (1987:10, 2006:45) points out, it can take about ten years to describe a never-before-studied language. Linguistic theories often change within that period of time. Of course, descriptive fieldwork without an underlying theory is impossible, but in descriptive fieldwork the theoretical approach itself should be descriptive and data-driven. Further comments about what a data-driven descriptive theoretical approach should look like are in Sections 11.4.2 and 12.1.

While there is no strong motivation for using non-descriptive theory-driven methodologies for fieldwork, such methods can be very helpful in developing specific fieldwork questions, as shown by Comrie (1988:5–6) and Rice (2006).

2.2.3.2 Comparative Theoretical Goals

There are three ways that languages can be compared: historically (including genetically), areally, and typologically.

The historical goals of fieldwork involve the collection of data so as to compare languages to determine genetic or other historical relationships. Grimes (1995:4–16), Vaux and Cooper (1999:165–180), and Vaux et al. (2007:351–381) are good sources of information on this. For most historical linguists, historically oriented fieldwork will first be the collection of basic vocabulary for the application of the comparative method.

Areal goals of fieldwork involve the collection of data useful for tracing mutual influences between languages, i. e. language contact. Four exemplary works on language contact based on extensive fieldwork are Haugen (1969) on Norwegian–American English contact; Hill and Hill (1986) on Nahuatl (Mexicano)-Spanish contact; Bakker (1997) on Mitchif, a mixed Cree–French language of Canada; and Aikhenvald (2002) on language contact in the Vaupes area of Amazonia. Older literature and references are in Weinreich (1974).

---

[8] What we call "theory" in this section is generally called "formal linguistic theory". The problem with the term "formal linguistic theory" is that it is understood to apply primarily to the Chomskyan paradigm, glossing over the fact that some functionalist theories are just as non-descriptive as Chomskyan formal linguistics.

Typological goals of fieldwork involve collecting data useful for identifying language universals (Abbi 2001, Evans and Dench 2006:5) or language particulars, also called *rara* (Ladefoged and Everett 1996; Everett 2004, 2005). The literature on typology is vast; an extended discussion of sources and surveys for language typology is provided in Chapter 11.[9]

Baker (2005) distinguishes three views of typology in linguistics. In the generative or Chomskyan approach, only a few languages are compared, and therefore little fieldwork is required. In conventional typological studies, hundreds of languages are compared, albeit somewhat superficially, and the amount of fieldwork conducted per language varies considerably. An exemplary and prominent example of this type is Haspelmath et al. (2005). Baker advocates a "middle" way of doing typology which involves comparing ten or so languages, and carrying out a very substantial amount of fieldwork on each of them. It should be noted that this middle way is the way that linguistic typology was carried out by fieldworkers such as Boas, Sapir and Bloomfield (Section 3.1) The goal of fieldworkers should be, in our opinion, to carry out fieldwork that can feed into both Baker's "middle" way and the conventional way of carrying out typological studies.

2.2.3.3 Dialectological or Sociolinguistic Goals

There are two basic schools in the study of intralinguistic variation: the dialectological school, focusing on regional variation (Pickford 1956; Chambers and Trudgill 1980) and the sociolinguistic school, focusing on social variation (Labov 1972, 1984).[10] Should dialectological or sociolinguistic research be regarded as fieldwork? Lounsbury (1953:413–14) says yes: dialectological research is linguistic fieldwork. Munro (2003:130) says no: sociolinguistic research is not fieldwork. As we see it, both of these schools, regardless of ideological differences, use descriptive fieldwork techniques, and have written more extensively about them than descriptive fieldworkers. A survey of dialectological fieldwork is in Francis (1983). A good survey of sociolinguistic techniques is Milroy (1987). See also Section 12.2 for further references to sociolinguistic techniques.

Dialectological or sociolinguistic fieldwork goals are emphasized in some recent accounts of fieldwork on Romance languages; for example, López Morales (1994) for Spanish, focusing on dialectology and sociolinguistics; and Blanchet (2000) for French, taking an ethno-sociolinguistic approach.[11]

---

[9] Typological fieldwork is also important from a terminological point of view, since the terminology used in documentary and descriptive fieldwork is based on typological findings, whereas the terminology for historical and areal fieldwork can be more easily constrained to those fields.

[10] Two recent discussions of fieldwork by Vaux and Cooper (1999:149–164) and by Vaux et al. (2007:315–349), treat issues of dialectological and sociolinguistic fieldwork together.

[11] Blanchet (2000) is interesting in that it covers both method and theory. However, the methodological part of Blanchet (2000) is also quite theoretical, and gives little practical advice.

Related to dialectological and sociolinguist goals is the issue of determining mutual intelligibility among related varieties, or the measurement of dialect distance. SIL linguists have recently been preoccupied with mutual intelligibility testing for practical reasons. Indeed, it is connected with the question of how many language varieties the Bible needs to be translated into. The fundamental work is Casad (1974), and the most recent account on this topic is Grimes (1995). Older discussions include Voegelin and Harris (1951), Hickerson et al. 1952, Smalley (1957), and Wolff (1959).

2.2.3.4 Goals Regarding the Study of Language, Culture, and Cognition

Some fieldworkers, mostly but not uniquely linguistic anthropologists, will be interested in the issue of the relationship between culture and language, i.e. does language condition culture, or vice-versa, or both. Similarly, they will ask whether language conditions cognition, or vice-versa, or both. These relationships are best exemplified in Lucy (1985, 1992a, b), Gumperz and Levinson (1996), Enfield (2002), and Everett (2005).

2.3 Aspirations and Limitations of Linguistic Fieldworkers

To conclude our chapter on the goals of fieldwork, we consider the personal aspirations of the fieldworker. First, who does the fieldworker want to be or become by conducting fieldwork? The field linguist wants to be more than an amiable and flashy character with a fancy hat like Indiana Jones (Bowern 2008:13–14). Nor does s/he want to be a nerdy character fidgeting on an uncomfortable bench with a fancy laptop which acts as a metaphorical wall between him/her and the puzzled speaker. The fieldworker might like working alone, but may also want to avoid the negative stereotype of the "Lone Ranger linguist", labeled as such by Dwyer (2006:54) as a caricature of the go-it-alone colonial fieldworker.[12] Perhaps the field linguist has humanitarian aspirations and would like to assume a personality similar to those of members of organizations like Doctors without Borders. Aren't field linguists ultimately "Linguists Without Borders"? They come in, sometimes live with the people for a while, and do good work, and maybe even help to save a language from extinction. The educational and humanitarian goals of training native speakers for language preservation, or of raising the profile of a language and its speakers are certainly fulfilling. All these characterizations of the field linguist exist and typically the individual finds himself/herself negotiating between several personae. In any case, linguistic fieldwork is intellectually exciting, as described in Abbi (2001), Bowern (2008), Crowley (2007), Aikhenvald (2007:4, 9), and the articles of Newman

[12] Australianists call this caricature the "Crocodile Dundee Fieldwork Model", as in the following blog by Jane Simpson: http://blogs.usyd.edu.au/elac/2007/04/theres_fieldwork_and_theres_fi.html

and Ratliff (2001), and personally fulfilling. Fieldworkers get to meet new people, and regardless of whether or not they visit exotic places, they create something new, or reveal something new to the world (Abbi 2001; Dixon 2007).

We would like to finish this chapter by adding three roles to characterize a descriptive field linguist, limitations and all. We will call these comparisons: the field linguist as astronomer, the field linguist as textual critic, and the field linguist as piano tuner.

Field linguists are like astronomers. Astronomy is a science where observations are paramount. Astronomers cannot travel to the stars and planets of outer space to see what they are really like, and they have to rely on whatever they can observe, at a distance of many light-years. The same thing is true, *mutatis mutandis*, with linguistic fieldwork. Field linguists cannot get into a speaker's brain and see which neuron does what when a particular grammatical construction is used (assuming, with Chomsky, that there is a language organ in there somewhere). All they can do is observe what comes out of the speaker's mouth. If an astronomer observes and describes a black hole or quasar or whatever in a part of the universe, regardless of whether it fits into someone's theory or not, s/he can publish that observation in a scientific journal. Like astronomers, field linguists have to observe and describe linguistic facts regardless of whether they fit into someone's theory or not, and hopefully they can publish their findings as well.

Field linguists are also like textual critics. As with the methodology of textual criticism, it is not possible to describe fieldwork methodologies in a totally explicit way. Indeed, fieldwork is never mechanical; intuition is at work, and it is as much an art as a science to do good fieldwork. Metzger (1992:219), who was for years the dean of New Testament Greek textual criticism in the United States, quotes an essay by the textual critic A. E. Housman as follows:

> A textual critic engaged upon his business is not at all like Newton investigating the motions of the planets: he is much more like a dog hunting for fleas. If a dog hunted for fleas on mathematical principles, basing his researches on statistics of area and population, he would never catch a flea except by accident. They require to be treated as individuals; and every problem which presents itself to the textual critic must be regarded as possibly unique.

Certainly, the fieldworker hopes that most problems s/he encounters will not be unique, but s/he must be prepared for that possibility.

Finally, and maybe most surprisingly, field linguists are also like piano tuners. If you have a piano, you must have it tuned occasionally. You will notice that piano tuners come in two versions: most bring equipment to calibrate the pitch of each key, but some bring no equipment: they have perfect pitch, and tune the piano entirely by ear. We tend to put more trust in the piano tuner who brings equipment, but on the other hand, we would not like a piano tuner who has no ear for pitch at all. In the same way, we expect the fieldworker to bring some equipment to the field, but at the same time we should look dimly upon a fieldworker who has to rely entirely on pitch tracking equipment to figure out what tones the language has and lexicographic software to determine the shape of a dictionary.

References

Abbi, Anvita. 2001. *A Manual of Linguistic Field Work and Indian Language Structures.* (Lincom Handbooks in Linguistics 17.) Munich: Lincom Europa.

Abley, Mark. 2003. *Spoken here: travels among threatened languages*. New York: Mariner books.

Aikhenvald, Alexandra Y. 2002. *Language Contact In Amazonia*. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y. 2007. Linguistic fieldwork: setting the scene. *Sprachtypologie und Universalienforschung* (Focus on Linguistic Fieldwork, ed. by Alexandra Y. Aikhenvald) 60(1):3–11.

Ameka, Felix K., Alan Dench, and Nicholas Evans, eds. 2006. *Catching language: The Standing Challenge of Grammar Writing*. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Amery, Rob, and Mary-Anne Gale. 2008. But our language was just asleep: a history of language revival in Australia. In *Encountering Aboriginal Languages: Studies in the History of Australian Linguistics*, ed. by William B. McGregor, 339–382. Canberra: Pacific Linguistics. Research School of Pacific and Asian Studies, The Australian National University.

Austin, Peter K., ed. 2003. *Language Documentation and Description*. Volume 1. London: School of Oriental and African Studies.

Austin, Peter K., ed. 2004. *Language Documentation and Description*. Volume 2. London: School of Oriental and African Studies.

Austin, Peter K., ed. 2005. *Language Documentation and Description*. Volume 3. London: School of Oriental and African Studies.

Austin, Peter K., ed. 2007. *Language Documentation and Description*. Volume 4. London: School of Oriental and African Studies.

Austin, Peter K., ed. 2008. *Language Documentation and Description*. Volume 5. London: School of Oriental and African Studies.

Austin, Peter K., ed. 2009. *Language Documentation and Description*. Volume 6. London: School of Oriental and African Studies.

Austin, Peter K., ed. 2010a. *Language Documentation and Description*. Volume 7. (Special issue: Lectures in Language Documentation and Description.) London: School of Oriental and African Studies.

Austin, Peter K. 2010b. Current Issues in Language Documentation. In *Language Documentation and Description*, Vol. 7, ed. by Peter K. Austin, 1–3. London: School of Oriental and African Studies.

Austin, Peter K., and Lenore Grenoble. 2007. Current Trends in Language Documentation. In *Language Documentation and Description*, Vol. 4, ed. by Peter K. Austin, 12–25. London: School of Oriental and African Studies.

Austin, Peter K., and Julia Sallabank, eds. 2010. *Handbook of Endangered Languages*. Cambridge: Cambridge University Press.

Austin, Peter K., and Andrew Simpson, eds. 2007. *Endangered languages*. Hamburg: Helmut Buske Verlag.

Baker, Mark C. 2005. The Impact of Typology on Theoretical Syntax and Morphology. Paper presented at the Typology Workshop of the Linguistic Society of America Meeting, January 9th, 2006. Oakland, CA.

Bakker, Peter. 1997. *A Language of Our Own. The Genesis of Michif, the Mixed Cree-French Language of the Canadian Metis*. (Oxford Studies in Anthropological Linguistics 10.) New York and Oxford: Oxford University Press.

Basset, A. 1951. L'enquête linguistique. *Conférence de l'Institut de Linguistique de l'Université de Paris. X*. Années 1950–1951, 7–22. Paris: Librairie C. Klincksieck.

Bausani, Alessandro. 1974. Nuovi materiali sulla lingua Chono. In *Atti del XL Congresso Internazionale degli Americanisti, Roma-Genova 3–10 Settembre 1972*. Vol. 3:107–116. Genova: Tilgher.

Beeler, Madison S. 1967. *The Ventureño Confesionario of José Señán, O.F.M*. (University of California Publications in Linguistics 47.) Berkeley and Los Angeles: University of California Press.

Beeler, Madison S. 1971. Noptinte Yokuts. In *Studies in American Indian Languages*, ed. by Jesse Sawyer, 11–76. (University of California Publications in Linguistics 65.) Berkeley and Los Angeles: University of California Press.

Blanchet, Philippe. 2000. *La linguistique de terrain. Méthode et théorie. Une approche ethno-sociolinguistique*. Rennes: Presses Universitaires de Rennes.

Bowern, Claire. 2008. *Linguistic Fieldwork. A Practical Guide*. New York: Palgrave MacMillan.

Bradley, David, and Maya Bradley, eds. 2002. *Language Endangerment and Language Maintenance*. London: RoutledgeCurzon.

Brenzinger, Matthias, ed. 1998. *Endangered Languages in Africa*. Cologne: Rüdiger Köppe.

Brenzinger, Matthias, ed. 2007. *Language Diversity Endangered*. Berlin and New York: Mouton de Gruyter.

Bunn, Gordon. 1974. *Golin Grammar*. (Workpapers in Papua New Guinea Languages 5.) Ukarumpa, Papua New Guinea: Summer Institute Linguistics.

Burnaby, Barbara, and Jon Reyhner, eds. 2002. *Indigenous Languages Across the Community*. Flagstaff: Northern Arizona University. Online: http://jan.ucc.nau.edy/~jar/ILAC/

Cantoni, Gina, ed. 1996. *Stabilizing Indigenous Languages*. Flagstaff, Arizona: Northern Arizona University. Online: http://www.ncela.gwu.edu/pubs/stabilize/index.htm, and http://jan.ucc.nau.edu/~jar/SIL_Appendix.html

Casad, Eugene. 1974. *Dialect Intelligibility Testing*. Dallas, TX: SIL.

Chambers, Jack K., and Peter Trudgill. 1980. *Dialectology*. (Cambridge Textbooks in Linguistics.) Cambridge: Cambridge University Press.

Chomsky, Noam A. 1957. *Syntactic Structures*. The Hague: Mouton.

Chomsky, Noam A. 1964. Current Issues in Linguistic Theory. In *The Structure of Language. Readings in the Philosophy of Language*, ed. by Jerry A. Fodor and Jerrold J. Katz, 50–118. Englewood Cliffs, NJ: Prentice-Hall.

Colinet, Ph. 1896. Het dialect van Aalst, eene phonetisch-historische studie. *Leuvense Bijdragen* 1.1–59, 97–206, 223–308.

Comrie, Bernard. 1988. The Role of the field linguist. *Notes on Linguistics* 41:4–6.

Comrie, Bernard. 2007. Documenting and/or Preserving Endangered Languages. In *The Vanishing Languages of the Pacific Rim*, Osahito Miyako, Osamu Sakiyama, and Michael Krauss, 25–34. Oxford: Oxford University Press.

Craig, Colette G. 1992. A constitutional response to language endangerment: The case of Nicaragua. *Language* 68(1):17–24.

Crowley, Terry. 2007. *Field Linguistics. A Beginner's Guide*. Oxford: Oxford University Press.

Crystal, David. 2000. *Language death*. Cambridge: Cambridge University Press.

Dalby, Andrew. 2003. *Language danger. The loss of linguistic diversity and the threat to our future*. New York: Columbia University Press.

De Dominicis, Amedeo, ed. 2006. *Undescribed and Endangered Languages; the Preservation of Linguistic Diversity*. Newcastle upon Tyne: Cambridge Scholars Publishing.

De Wolf, Paul P. 1997. *Esbozo del Mayo Sonorense*. Two vols. Hermosillo, Sonora: Editorial Unison, Universidad de Sonora.

DeChicchis, Joseph. 1995. The current state of Ainu language. *Journal of Multilingual and Multicultural Development* 16(1/2):103–24.

Dixon, R. M. W. 2007. Field Linguistics: a minor manual. *Sprachtypologie und Universalienforschung* (Focus on Linguistic Fieldwork, ed. by Alexandra Y. Aikhenvald) 60(1):12–31.

Dixon, R. M. W. 2010. *Basic Linguistic Theory. Volume 1 Methodology*. Oxford: Oxford University Press.

Dixon, R. M. W., and Alexandra Y. Aikhenvald, eds. 1999. *The Amazonian Languages*. (Cambridge Language Surveys.) Cambridge: Cambridge University Press.

Dwyer, Arienne. 2006. Ethics and practicalities of cooperative fieldwork and analysis. In *Essentials of Language Documentation*, ed. by Jost Gippert, Nikolaus P. Himmelmann, and

Ulrike Mosel, 31–66. (Trends in Linguistics. Studies and Monographs 178.) Berlin and New York: Mouton de Gruyter.

Enfield, Nicholas J., ed. 2002. *Ethnosyntax. Explorations in Grammar and Culture*. Oxford and New York: Oxford University Press.

England, Nora C. 1992. Doing Mayan linguistics in Guatemala. *Language* 68(1):29–35.

Evans, Nicholas. 2010. *Dying Words. Endangered Languages and What They Have to Tell Us*. Malden, MA: Wiley-Blackwell.

Evans, Nicholas, and Alan Dench. 2006. Introduction: Catching language. In *Catching language: The Standing Challenge of Grammar Writing*, ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 1–39. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Evans, Nicholas, and Stephen Levinson. 2009. The Myth of Language Universals: Language Diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32(5):429–492.

Everett, Daniel L. 2001. Monolingual field research. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 166–188. Cambridge: Cambridge University Press.

Everett, Daniel L. 2004. Coherent fieldwork. In *Linguistics today—Facing a greater challenge*, ed. by Piet van Sterkenburg, 141–162. Amsterdam and Philadelphia: John Benjamins.

Everett, Daniel L. 2005. Cultural Constraints on Grammar and Cognition in Pirahã. Another Look at the Design Features of Human language. *Current Anthropology* 46(4):621–646.

Fishman, Joshua, ed. 2001. *Can threatened languages be saved? Reversing language Shift, Revisited: A 21st Century perspective*. Clevedon: Multilingual Matters.

Foley, William A. 2002. Field methods. In *The Linguistics Encyclopedia*. Second ed., ed. by Kirsten Malmkjær, 131–137. London and New York: Routledge.

Francis, W. N. 1983. *Dialectology. An Introduction*. London and New York: Longman.

Gippert, Jost, Nikolaus P. Himmelmann, and Ulrike Mosel, eds. 2006. *Essentials of Language Documentation*. (Trends in Linguistics. Studies and Monographs 178.) Berlin and New York: Mouton de Gruyter.

Good, Jeff. 2006a. *The Descriptive Grammar as a (Meta)Database*. Online: http://emeld.org/workshop/2004/jcgood-paper.html

Good, Jeff. 2006b. *The Ecology of documentary and descriptive linguistics*. Position paper for the 2006 E-MELD Workshop on Digital Language Documentation: Tools and Standards: The State of the Art. Online: http://www.acsu.buffalo.edu/~jcgood/jcgood-Tool_Ecology.pdf

García, Bartolomé. 1760. *Manual para administrar los Santos Sacramentos de Penitencia, Eucharistia, Extrema-Uncion, y Matrimonio: Dar Gracias Despues de Comulgar, y Ayudar a Bien Morir A los Indios de las Naciones: Pajalates, Orejones, Pacaos, Pacóas, Tilijayas, Alasapas, Pausanes, y otras muchas diferentes, que se hallan en las Missiones del Rio de San Antonio, a Rio Grande, pertenecientes â el Colegio de la Santissima Cruz de la Cuidad de Queretaro, como son: Los Pacuâches, Mescâles, Pampôpas, Tâcames, Chayopînes, Venados, Pamâques, y toda la Juventud de Pihuiques, Borrados, Sanipaos, y Manos de Perro*. [Mexico City]: Imprenta de los Herederos de Dona Maria de Rivera.

Grenoble, Lenore, and Lindsay J. Whaley, eds. 1998. *Endangered languages: Language loss and community response*. Cambridge: Cambridge University Press.

Grenoble, Lenore, and Lindsay J. Whaley. 2006. *Saving languages: An introduction to language revitalization*. Cambridge: Cambridge University Press.

Grimes, Joseph E. 1995. *Language Survey Reference Guide.* Dallas: Summer Institute of linguistics.

Grinevald, Colette. 2007. Linguistic Fieldwork among Speakers of Endangered Languages. In *The Vanishing Languages of the Pacific Rim*, Osahito Miyako, Osamu Sakiyama, and Michael Krauss, 35–76. Oxford: Oxford University Press.

Gudschinsky, Sarah C. 1957. *Handbook of Literacy*, rev. ed. Santa Ana, CA: Summer Institute of Linguistics.

Gumperz, John. J. and Stephen C. Levinson. 1996. *Rethinking Linguistic Relativity*. (Studies in the Social and Cultural Foundations of Language 17.) Cambridge: Cambridge University Press.

Hale, Kenneth L. 1972. Some questions about anthropological linguistics: The role of native knowledge. In *Reinventing anthropology*, ed. by Dell Hymes, 382–97. New York: Pantheon Books.

Hale, Kenneth L. 1992a. On Endangered languages and the safeguarding of diversity. *Language* 68(1):1–3.

Hale, Kenneth L. 1992b. Language endangerment and the human value of linguistic diversity. *Language* 68(1):35–42.

Harrison, K. David. 2007. *When Languages Die*. *The Extinction of the World's Languages and the Erosion of Human Knowledge*. Oxford: Oxford University Press.

Harrison, K. David, David S. Rood, and Arienne Dwyer, eds. 2008. *Lessons from Documented Endangered Languages*. (Typological Studies in Language 78.) Amsterdam and Philadelphia: John Benjamins.

Haspelmath, Martin, Matthew S. Dryer, David Gil, and Bernard Comrie, eds. 2005. *The World Atlas of Language Structures*. Oxford and New York: Oxford University Press.

Haugen, Einar I. 1969. *The Norwegian Language in America. A Study in Bilingual Behavior*. Bloomington and London: Indiana University Press.

Hickerson, H., G. D. Turner, and Nancy P. Hickerson. 1952. Testing procedures for estimating transfer of information among Iroquois dialects and languages. *International Journal of American Linguistics* 18:1–8.

Hill, Jane, H., and Kenneth C. Hill. 1986. *Speaking Mexicano. Dynamics of a Syncretic language in Central Mexico*. Tucson: University of Arizona Press.

Himmelmann, Nikolaus P. 1998. Documentary and descriptive linguistics. *Linguistics* 36: 161–195. Online: http://corpus.linguistics.berkeley.edu/~ling240/himmelmann.pdf or http://hrelp.org/events/workshops/eldp2005/reading/himmelmann.pdf

Himmelmann, Nikolaus P. 2006. Language documentation: What is it and what is it good for? In *Essentials of Language Documentation*, ed. by Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel, 1–30. (Trends in Linguistics. Studies and Monographs 178.) Berlin and New York: Mouton de Gruyter.

Hinton, Leanne, and Kenneth Hale, eds. 2001. *The Green Book of Language Revitalization in Practice*. San Diego, CA: Elsevier Science.

Hinton, Leanne, Matt Vera, Nancy Steele, and the Advocates for Indigenous Language Survival. 2002. *How to Keep Your Language Alive: A Common Sense Approach to One-on-One Language Learning*. Berkeley, CA: Heyday Books.

Hyman, Larry M. 2001. Fieldwork as a state of mind. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 15–33. Cambridge: Cambridge University Press.

Janse, Mark, and Sijmen Tol, eds. 2003. *Language death and language maintenance: theoretical, practical and descriptive approaches.* Amsterdam and Philadelphia: John Benjamins.

Jeanne, Laverne Masayesva. 1992. An institutional response to language endangerment: A proposal for a Native American Language Center. *Language* 68(1):24–28.

Johnson, Heidi. 2004. Language documentation and archiving, or how to build a better corpus. In *Language Documentation and Description*. Volume 2, ed. by Peter K. Austin, 140–53. London: SOAS, Hans Rausing Endangered Languages Project.

Kasten, Erich, ed. 1998. *Bicultural Education in the North: Ways of Preserving and Enhancing Indigenous Peoples' Languages and Traditional Knowledge*. Münster and New York: Waxmann.

Kennedy, Graeme. 1998. *An Introduction to Corpus Linguistics*. London: Longman.

Krauss, Michael E. 1992. The world's languages in crisis. *Language* 68(1):4–10.

Labov, William. 1972. Some principles of linguistic methodology. *Language in Society* 1:97–120.

Labov, William. 1984. Field methods of the project on linguistic change and variation. In *Language in use. Readings in Sociolinguistics*, ed. by John Baugh and Joel Sherzer, 28–35. Englewood Cliffs, N.J.: Prentice-Hall.

Ladefoged, Peter. 1992. Another view of endangered languages. *Language* 68(4):809–11.

Ladefoged, Peter, and Daniel Everett. 1996. The status of phonetic rarities. *Language* 72(4): 794–800.

Lehmann, Christian. 1999. Documentation of endangered languages. A priority task for linguistics. *ASSIDUE, Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt* Nr. 1. Erfurt: Seminar für Sprachwissenschaft, Philosophische Fakultät, Universität Erfurt.

Lehmann, Christian. 2004. Documentation of grammar. In *Lectures on endangered languages 4. Papers from the Kyoto Conference on Language Endangerment, 2001*, ed. by Osamu Sakiyama, Fubito Endo, Honore Watanabe, and Fumiko Sasama, 61–74. Osaka: Osaka Gakuin University. (Endangered Languages of the Pacific Rim Publication Series, C–004)

Lindenfeld, Jacqueline. 1973. *Yaqui Syntax*. (University of California Publications in Linguistics 76.) Berkeley and Los Angeles: University of California Press.

López Morales, Humberto. 1994. *Métodos de Investigación Lingüística*. Salamanca: Ediciones Colegio de España.

Lounsbury, Floyd. 1953. Field methods and techniques in linguistics. In *Anthropology today: an encyclopedic inventory*, ed. by Alfred L. Kroeber, 401–416. Chicago: University of Chicago Press.

Lucy, John A. 1985. Whorf's View of the Linguistic Mediation of Thought. In *Semiotic Mediation*, ed. by Elizabeth Mertz and Richard J. Parmentier. Orlando: Academic Press Inc.

Lucy, John A. 1992a. *Language, Diversity and Thought: A Reformulation of the Linguistic Relativity Hypothesis*. Cambridge: Cambridge University Press.

Lucy, John A. 1992b. *Grammatical Categories and Cognition: A Case Study of the Linguistic Relativity Hypothesis*. Cambridge: Cambridge University Press.

Machoni de Cerdeña, Antonio. 1877. *Arte y vocabulario de la lengua lule y tonocoté*. Buenos Aires: Pablo E. Coni.

MacKay, Carolyn J., and Frank R. Trechsel. 2005. *Totonaco de Misantla, Veracruz*. (Archivo de lenguas indígenas de México.) Mexico City: El Colegio de México.

Matsumura, Kazuto, ed. 1998. *Studies in endangered languages: Papers from the international symposium on endangered languages*, Tokyo, November 18–20, 1995. (International Clearinghouse for Endangered Languages linguistic studies 1.) Tokyo: Hituzi Syobo.

Matthews, G. Hubert. 1965. *Hidatsa Syntax*. (Papers on Formal Linguistics 3.) The Hague: Mouton.

McEnery, Tony, and Andrew Wilson. 1996. *Corpus Linguistics*. Edinburgh: Edinburgh University Press.

Meijs, Willem, ed. 1987. *Corpus Linguistics and Beyond*. Amsterdam: Rodopi.

Metzger, Bruce M. 1992. *The Text of the New Testament. Its Transmission, Corruption, and Restoration*. Third enlarged ed. New York and Oxford: Oxford University Press.

Milroy, Lesley. 1987. *Observing and Analysing Natural Language. A Critical Account of Sociolinguistic method*. (Language in Society 12.) Oxford: Basil Blackwell.

Miyaoka, Osahito, Osamu Sakiyama, and Michael Krauss, eds. 2007. *The Vanishing Languages of the Pacific Rim*. Oxford: Oxford University Press.

Mosel, Ulrike. 1987. *Inhalt und Aufbau deskriptiver Grammatiken (How to write a grammar)*. Institut für Sprachwissenschaft der Universität zu Köln, Arbeitspapier Nr. 4 (Neue Folge.) Cologne: Institut für Sprachwissenschaft, Universität Köln.

Mosel, Ulrike. 2001. Linguistic Fieldwork. In *International Encyclopedia of the Social and Behavioral Sciences*, Vol. 13, ed. by Niel J. Smelser and Paul B. Baties, 8906–8910.

Mosel, Ulrike. 2006. Grammaticography: The art and craft of writing grammars. In *Catching language: The Standing Challenge of Grammar Writing*, ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 11–68. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Moseley, Christopher. 2007. *Encyclopedia of the World's Endangered Languages*. Oxford and New York: Routledge.

Mufwene, Salikoko S. 1998. The Ecology of Language: New Imperatives in Linguistics Curricula. *Studies in the Linguistic Sciences* 28(2):135–145.

Munro, Pamela. 2003. Field Linguistics. In *The Handbook of Linguistics*, ed. by Mark Aronoff and Janie Rees-Miller, 130–149. Malden, MA: Blackwell Publishing.

Nathan, David. 2006. Thick interfaces: Mobilizing language documentation with multimedia. In *Essentials of Language Documentation*, ed. by Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel, 363–379. (Trends in Linguistics. Studies and Monographs 178.) Berlin and New York: Mouton de Gruyter.

Nettle, Daniel. 1998. *Linguistic diversity*. New York: Oxford University Press.

Nettle, Daniel, and Suzanne Romaine. 2000. *Vanishing voices: the extinction of the world's languages*. New York: Oxford University Press.

Newman, Paul. 1998. We have seen the enemy and it is us: the endangered languages issue as a hopeless cause. *Studies in the Linguistic Sciences* 28(2):11–20.

Newman, Paul, and Martha Ratliff, eds. 2001. *Linguistic fieldwork*. Cambridge: Cambridge University Press.

Oostdijk, Nelleke. 1988. Corpustaalkunde in perspectief. *TTT, Interdisciplinair Tijdschrift voor Taal en Tekstwetenschap* 8(3):209–224.

Ostler, Nicholas, ed. 1998. *Endangered Languages: What Role for the Specialist? Proceedings of the second FEL conference*. Bath: The Foundation for Endangered Languages.

Pauwels, J. L. 1958. *Het Dialect van Aarschot en Omstreken*. Two vols. Tongeren: Belgisch Interuniversitair Centrum voor Neerlandistiek.

Payne, Thomas E, and David J. Weber, eds. 2007. *Perspectives on Grammar Writing*. (Benjamins Current Topics 11.) Amsterdam and Philadelphia: John Benjamins. [originally published as 2006, Special Issue of *Studies in Language* 30(2)]

Pickford, G. R. 1956. American linguistic geography: a sociological appraisal. *Word* 12:211–33.

Pittman, Dean. 1948. *Practical Linguistics. A Textbook and Field Manual of Missionary Linguistics*. Cleveland, OH: Mid-Missions.

Reyhner, Jon, ed. 1997. *Teaching Indigenous Languages*. Flagstaff: Northern Arizona University. Online: http://jan.ucc.nau.edu/~jar/TIL_Contents.html

Rice, Keren. 2006. Let the language tell its story? The role of linguistic theory in writing grammars. In *Catching language: The Standing Challenge of Grammar Writing*, ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 235–268. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Robins, Robert H., and Eugenius M. Uhlenbeck, eds. 1991. *Endangered Languages*. Oxford and New York: Berg Publishers (Diogenes Library).

Ruz, Mario Humberto, and Diana Birrichaga, eds. 1997. *Las lenguas del Chiapas Colonial. Manuscritos en la biblioteca nacional de París*. Vol. dos: lengua zoque. Mexico City: Universidad Nacional Autónoma de México.

Sakiyama, Osamu, ed. 2004. *Lectures on Endangered Languages 5*. (Endangered Languages of the Pacific Rim C-005.) Osaka: Osaka Gakuin University.

Sakiyama, Osamu, and Fubito Endo, eds. 2001. *Lectures on Endangered Languages 2*. (Endangered Languages of the Pacific Rim C-002.) Osaka: Osaka Gakuin University. Online: elpr@utc.osaka-gu.ac.jp

Sakiyama, Osamu, Fubito Endo, Honoré Watanabe, and Fumiko Sasama, eds. 2004. *Lectures on Endangered Languages 5*. (Endangered Languages of the Pacific Rim C-004.) Osaka: Osaka Gakuin University.

Samarin, William J. 1967. *Field Linguistics. A Guide to Linguistic Field Work*. New York: Holt, Rinehart and Winston.

Sampson, Geoffrey. 2002. *Empirical Linguistics*. London and New York: Continuum. FI

Sapir, Edward. 1933. La réalité psychologique des phonèmes. *Journal de Psychologie Normale et Pathologique* 30:247–65.

Seay, Elizabeth. 2003. *Searching for Lost City*. On the Trail of America's Native Languages. Guiford, Conn.: The Lyons Press.

Seifart, Frank. 2006. Orthography development. In *Essentials of Language Documentation*, ed. by Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel, 275–299. (Trends in Linguistics. Studies and Monographs 178.) Berlin and New York: Mouton de Gruyter.

Smalley, William A. 1957. Finding out how close related dialects are. Part II: conducting a dialect survey. *The Bible Translator* 8:114–26.

Teubert, Wolfgang, and Anna Cermáková. 2007. *Corpus Linguistics. A Short Introduction*. London and New York: Continuum.
Troike, Rudolph C. 1996. Sketch of Coahuilteco, a Language Isolate of Texas. In *Handbook of North American Indians*, Vol. 17, Languages, ed. by Ives Goddard, 644–665. Washington, D.C.: Smithsonian Institution.
Tuttle, Siri G. 2003. Archival Phonetics: Tone and Stress in Tanana Athabaskan. *Anthropological Linguistics* 45(3):316–36.
Vanacker, F. 1948. *Syntaxis van het Aalsters Dialect*. (Werken Uitgegeven door de Koninklijke Commissie voor Toponymie en Dialectologie 4) Tongeren: George Michiels.
Vaux, Bert, and Justin Cooper. 1999. *Introduction to Linguistic Field Methods*. (Lincom Course books in Linguistics 1.) Munich: Lincom Europa.
Vaux, Bert, Justin Cooper, and Emily Tucker. 2007. *Linguistic Field Methods*. Eugene, OR: Wipf & Stock.
Voegelin, Carl F., and Zellig S. Harris. 1951. Methods for Determining Intelligibility Among Dialects of Natural Languages. *Proceedings of the American Philosophical Society* 45:322–329.
Watahomigie, Lucille J., and Akira Y. Yamamoto. 1992. Local Reactions to perceived language decline. *Language* 68(1):10–17.
Weinreich, Uriel. 1974. *Languages in Contact. Findings and Problems*. The Hague & Paris: Mouton.
Wolff, Hans. 1959. Intelligibility and Inter-Ethnic Attitudes. *Anthropological Linguistics* 1(3):34–41.
Woodbury, Anthony C. 2003. Defining documentary linguistics. In *Language Documentation and Description,* Vol. 1, ed. by Peter K. Austin, 35–51. London: School of Oriental and African Studies.
Woodbury, Anthony C. 2007. On thick translation in linguistic documentation. In *Language Documentation and Description,* Vol. 4, ed. by Peter K. Austin, 120–135. London: School of Oriental and African Studies.
Wynne, Martin, ed. 2005. *Developing Linguistic Corpora: a Guide to Good practice*. Oxford: Oxbow Books. Online: http://ahds.ac.uk/linguistic-corpora/
Zaefferer, Dietmar. 2006. Realizing Humboldt's dream: Cross-linguistic grammatography as database creation. In *Catching language: The Standing Challenge of Grammer Wrinting*, ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 113–35. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Chapter 3
The History of Linguistic Fieldwork

3.1 Introduction

A full-fledged history of linguistic fieldwork would be an interesting subtopic within the history of linguistics, and would also be relevant to the history of ethnography, the history of European colonization, and the history of Christian missions. Such a study, which could be conceived as a history of human curiosity about other languages, still needs to be written. In this chapter, we will focus on the history of fieldwork insofar as it is relevant to the modern fieldworker.

Very little has been written on the history of linguistic fieldwork. The single most useful volume relating to this topic is McGregor (2008a), which contains an exemplary introduction on the history of research on Australian languages (McGregor 2008b). Even more significantly, to the best of our knowledge, it contains the only paper-length piece of literature whose title includes the term "history of fieldwork" specifically referring to linguistic fieldwork (McGregor 2008c). This pioneering paper, "History of fieldwork on Kimberley languages", is the first one to treat in detail the different methods and technologies used over the years by fieldworkers, though it only covers one geographical area of aboriginal Australia. This lone paper contrasts with the numerous historical accounts of anthropological or archeological fieldwork in the literature. Certainly, linguistic fieldworkers have been extraordinarily shy or suspiciously modest about their past methods and technologies!

In the remainder of this section, we will ask why the modern fieldworker might want to know something about the history of fieldwork (Section 3.1.1), and what the limitations on the study of this history are (Section 3.1.2). We will also briefly review the history of fieldwork funded by museums, universities and granting agencies, and conducted according to various European and American structuralist traditions (Section 3.1.3).

Then, in subsequent sections, we will discuss five types or characteristics of fieldwork that have existed at different historical periods: Christian missionary fieldwork (Section 3.2); fieldwork of what we will call the "gentleman-scholar" type (Section 3.3); fieldwork in less than optimal circumstances (Section 3.4);

S.L. Chelliah and W.J. de Reuse, *Handbook of Descriptive Linguistic Fieldwork*,
DOI 10.1007/978-90-481-9026-3_3,

situations where consultants were collaborators, rather than just subjects in fieldwork (Section 3.5); and fieldwork in academic traditions that have been less than supportive of, or conducive to fieldwork, such as the neo-grammarian school of the nineteenth century, or the Chomskyan tradition starting in the mid twentieth century (Section 3.6). These five types/characteristics of fieldwork overlap with each other topically and diachronically to various extents, particularly in the context of European colonialism. However, different conclusions can be drawn and different lessons learned by the modern fieldworker from this historical perspective. These conclusions and lessons are discussed in each section, and then for the chapter as a whole.

3.1.1 Why Study the History of Fieldwork?

There are two reasons why the modern fieldworker should be interested in the history of fieldwork. The first has to do with changing ethical standards involving human research. The fieldworker should learn what mistakes and errors were made in the past – these are more numerous, diverse, and imaginative than we care to believe – and learn to avoid them. The second relates to learning from the findings of previous researchers.

In the view of at least some of the practitioners of field linguistics, linguistics should be an inductive and experimental science. However, it is important to remind ourselves of the ethical problems occurring when an experimental science is practiced with human subjects. The history of fieldwork allows us to see how these ethical problems were dealt with, or ignored, in the distant and recent past.

We feel far removed from what was probably the first - and we cannot help hoping, mythical - linguistic experiment in the world. The Greek historian Herodotus, writing in the fifth century BCE, tells about the Egyptian king Psammetichus who conducted an experiment to determine which living language was the oldest. He had two newborns raised out of contact with human speech in order to find out the first word they would say after the stage of babbling. That word appeared to be “bekos”, which according to the king’s experts was the Phrygian word for “bread”. Therefore Phrygian, an extant language of Asia Minor, was considered to be the oldest language (Pedersen 1931/1962:3).

In recent times an experiment was suggested for scientifically studying the development of pidgins by arranging for a group of monolingual speakers, who did not speak each other’s languages, to live together on a desert island. Predictably, the experiment was rejected as unethical for a variety of reasons, including the fact that there would have been no way of satisfactorily explaining to the subjects the reason for their being on the island without defeating the purpose of the experiment (Crowley 2007:32–33). What this proposal shows is that the temptation for unethical experiments, as perhaps indulged in by Psammetichus, is still with us.

Because in the past individual fieldworkers were typically left alone setting their own ethical standards and regulating their own behavior, some early fieldworkers exhibited varied degrees of bad behavior towards speakers. So did many colonial administrators; missionaries; aid, social and medical workers; art collectors; and anthropologists (Kummer 1981). The study of the history of fieldwork thus helps us understand why potential linguistic consultants, in many areas of the world, have negative perceptions of outside researchers. Present-day ethical standards are discussed in Chapter 6.

The second reason why it is of interest and value to consider the history of fieldwork is that it provides an opportunity to learn from the findings of earlier generations of fieldworkers, particularly in the case of endangered language documentation. The fieldworker should learn to read, interpret, and understand older documentation, such as older grammars or field notes. Such older documentation, sometimes dismissed by modern fieldworkers as poor in quality and outdated, should be studied philologically, since a philological approach can determine precisely what is outdated and erroneous, and what data are relevant and useful to the modern linguist. This philological approach to archival data or grammars is discussed further in Section 5.2. It is obvious that a philological approach requires an understanding of the historical context in which the documents were composed. Therefore, extensive studies of the history of documentation of particular languages, of which the best recent example is Krauss (2006), are strong desiderata for the fieldworker.

It is true, of course, that there is no possibility of studying the history of fieldwork in an area where no fieldwork has previously been conducted. But it is overwhelmingly the case that most areas of the world have been visited by researchers who have either left an impression on the native people of that area, or have left behind some sort of document, however trivial or misconceived, on the language. History helps us understand and put into proper perspective these perceptions and/or documents.

3.1.2 Limitations on the Study of the History of Fieldwork

There are some limits to how well the history of linguistic fieldwork can be studied, since our knowledge of actual fieldwork techniques is very indirect. Until the twentieth century, linguists did not reflect on how fieldwork should be done; on the basis of available documents, we can only guess at the methods, techniques, procedures, and informant relations that existed. For example, from examining Busbecq's sixteenth century Crimean Gothic word list (see also Section 3.3.1), and the knowledge that it was elicited at a dinner, we can conclude that Busbecq must have been pointing at items visible above the table, such as parts of the human face, or bread, wine, and water, but not at items not visible in the context, such as feet (Rousseau 1991). In addition, motivations for collecting language data have varied considerably. Some people were simply curious or interested enough to collect samples of languages; others, such as colonial administrators or missionaries, collected language samples for purely practical reasons.

3.1.3 Fieldwork as a Funded Enterprise in European and American Traditions

Fieldwork funded by museums, universities or granting agencies, in various European structuralist traditions and in the American Boas-Sapir-Bloomfield tradition is relatively better documented than any other linguistic fieldwork. Therefore, to avoid redundancy, it will not be treated in detail in this chapter, although a synthesis of this sort of fieldwork from the point of view of the history of linguistic fieldwork still needs to be written. Basic information on the practices of early museum workers and fieldworkers such as John Rupert Firth (1890–1960), Franz Boas (1858–1942), Edward Sapir (1884–1939), Leonard Bloomfield (1887–1949), and their students, can be found in the references in the next paragraphs.

Materials for the study of American structuralism (more accurately called "descriptivist" (Sampson 1980:57–80), or "descriptive structuralism" (Hymes 2003:123) have been gathered by historians of linguistics and anthropology such as Hymes and Fought (1981). Discussions of Firth are in Firth et al. (1957). Discussions of Boas[1] are in Boas Yampolsky (1958), Canger (1994:1219), Darnell (1969, 1999), Mead and Bunzel (1960:461), Stocking (1992), Wax (1971:33–36), White (1963), and Campbell (1997:62–660). Discussions of Sapir are in Haas (1953), Cowan et al. (1986), Darnell (1990), Campbell (1997:69–72), and Mandeville and Scollon (2009:230–234). Discussions of Bloomfield are in Bloomfield (1962:vii), Fries (1961), Voegelin (1959c, 1960), Goddard (1987), and Campbell (1997:76–77). Further discussion of Boas, Sapir and Bloomfield is in Haas (1976) and further discussion of the resulting Americanist tradition is in Hymes (1976), Valentine and Darnell (1999), Mithun (1996), and Hill (2006).[2]

Materials for a history of American museum linguistic fieldwork are in Sturtevant and Stanley (1973), Goddard (1973), Landar (1976a:99–100), Goddard (1996), and Campbell (1997:57–62, 77–78). The eccentric John Peabody Harrington (1884–1961), who was an employee of the Smithsonian Institution, must be given special mention, as he was certainly the most productive and compulsive fieldworker ever. Accounts of Harrington's life and fieldwork are in Laird (1975), a biography by his ex-wife, and in Callaghan (1991), Golla (1991), Harrington (1989), Hinton (1994:195–209), James (1984), and Klar (1991).

Materials for a history of the Moscow school are in Kibrik (2007) and in Schulze (2005:331–343), and for the London school in Collins and Mees (1999:160–161), Firth et al. (1957), Newmeyer (1986:57–59), Sampson (1980:214), and Moore (2008:289–290).

---

[1] Boas was famous for insisting that the fieldworker learn as much of the language as possible. However, Boas' most famous student, the anthropologist Margaret Mead (1901–1978), had little interest in intensive language learning, and discussed her point of view in Mead (1939), and hinted at it in several biographical works (Mead 1972:139–140, 1977:30).

[2] References to obituaries and reminiscences on Boas, Sapir, and Bloomfield are listed in Landar (1976a:103).

3.2 The History of Christian Missionary Linguistic Fieldwork

3.2.1 Early Activities in Africa and Asia

Christian missionary interests played a large part during early Portuguese and Spanish colonization, which began in the fifteenth century. In contrast to other important religions such as Hinduism and Judaism, which were not interested in expanding through missionary efforts, and Islam, which required the exclusive use of Arabic as the language of proselytization and worship, Christianity has insisted on communication in the language of the people for missionary purposes (Wonderly and Nida 1963:105). As a result, the study of unwritten languages was almost exclusively carried out by Christian missionaries until the late eighteenth century (Hymes 1963:65).

The earliest text composed by a missionary in an African language was a catechism in the Kongo language (1624) and in 1659, a Kongo grammar by the Capuchin missionary Giacinto Brusciotto appeared. This was also the first grammar of any Bantu language, and while based on Latin grammar, the author was perceptive enough to discover and describe the nominal class system of Bantu (Gregersen 1977:93–9, Jungraithmayr and Möhlig 1983:56).

While China has its own venerable tradition of lexicography, the writing of grammars of Sinitic languages started with the arrival of missionaries, who wrote the first descriptions in the seventeenth century, with a focus on vernacular dialects, and who designed the earliest romanization systems. Of course, the local literati helped with the traditional spelling system. The earliest of these grammars were written by Spanish-speaking missionaries, who used as a model the 1481 Latin grammar by Elio Antonio de Nebrija (1444–1522) (Chappell 2006:441–442). This is the same Nebrija who later wrote an even more influential grammar of Spanish (see Section 3.2.2).

Eight grammars of Philippine languages were written by Spanish missionaries in the seventeenth and eighteenth centuries (Ridruejo 2004). García-Medall (2004) argues that the design of dictionaries of Philipino languages was influenced by the work of Spanish missionaries working on Mesoamerican languages.

The French Jesuit Alexandre de Rhodes (1591–1660) arrived in Vietnam in 1620. He wrote the first Vietnamese catechism, and created the Roman spelling system of Vietnamese (Gregerson 1981).

Portuguese Jesuits studied Japanese in the sixteenth and seventeenth centuries, and, as in China, their studies were greatly helped by the existence of a written tradition (Maruyama 2004). In Asia and northern Africa, most missionaries preferred working with languages that had existing writing systems, rather than working only with the spoken language, as they had to do in the Americas. In fact, early Western scholarship on written languages such as Arabic, Syriac, Persian, Coptic, Ethiopian, Sanskrit, Tamil, Chinese, Tibetan, Burmese, Thai, Malay, and Japanese (carried out, of course, with the help of local literati) was more highly regarded than the missionary documentation of spoken varieties (Firth 1937 [1964]:57–60).

The first European to make systematic observations of spoken Indian languages was the Jesuit Thomas Stephens, who lived in Portuguese India between 1579 and 1619. Foreshadowing the well-known pronouncements of Sir William Jones (1786) on the affinities of Sanskrit to European languages, Stephens wrote in 1583, "Many are the languages of these places… Their pronunciation is not disagreeable and their structure is allied to Greek and Latin. The phrases and constructions are of a wonderful kind." (quoted by Firth 1937:57).

3.2.2 Missionary Beginnings in Latin America

Linguistic fieldwork in the New World started soon after the conquest of Mexico by Cortez. The Franciscan missionary Fray Pedro de Gante arrived in 1523 and had compiled a *doctrina* ("catechism") in Nahuatl by 1553. Twelve other Franciscan friars arrived in 1524. With the assistance of Alonso de Molina, the child of a Spanish widow who had learned Nahuatl from his native playmates, they wrote more than 80 *doctrinas*, *confesionarios* ("confessionals"), *sermonarios* ("collections of sermons"), *artes* ("grammars"), *vocabularios* ("vocabularies") and scriptural translations in Nahuatl, the major language of the Aztec realm.

The Dominicans arrived in 1526 and the Augustinians in 1533. They studied and preached in Nahuatl as well in more than a hundred other languages of New Spain. Between 1524 and 1572, a total of 109 known works in or on these languages was produced. The first product of the Mexico City printing press was a *doctrina* in Nahuatl, which came out in 1539 (Ostler 2004:39). Some friars were apparently in command of more than one language; for example Fray Andres de Olmos wrote an *arte* for Nahuatl and is alleged to have written *artes* for Huastec and Totonac as well (McQuown 1976:105–106, Suárez 1983:2).

It also appears that every order of missionaries wanted to have its own Nahuatl grammar, and that the Jesuit missionaries were more perceptive phonetically than the Franciscan or the Augustinian missionaries in the difficult matters of writing vowel length and the glottal stop (Canger 1990:107–110; 1997).

Since there were no written languages in New Spain at the time (with the exception of hieroglyphic Mayan, which was not studied and which only came to be understood much later), preparing descriptions of these languages must have been the result of a massive fieldwork enterprise. This enterprise was considered a natural necessity for conversion and Christianization. The model for missionaries' grammars was Elio Antonio de Nebrija's (or Lebrija's) grammar of Spanish, which first appeared in 1492 and which was revolutionary in being the first full grammar of any European vernacular (Rowe 1974:361, Zimmermann 1997b:10–11, Ostler 2005:365, Chappell 2006:441). It systematically compared the grammar of contemporary Spanish (not known as a native language by all friars) to that of Latin (known by all friars), and continued as a model for grammatical description up to 1821. Some of the grammars were particularly perceptive. For example, Carochi's grammar of Nahuatl, published in 1645 (Canger 1997), included accurate phonetic detail for which Nebrija's work could not have been the model (McQuown 1976:108–109; Suárez 1983:4).

In the Andean region, the first grammar and lexicon of Quechua was written in 1560 by the Dominican Domingo de Santo Tomás (ca. 1499–1570). Like all his contemporaries, Santo Tomás was influenced by Nebrija and by Latin grammar, but he did point out that there were features of Quechua which were not found in Latin or Spanish. The most perceptive early grammar and dictionary of Quechua was written in 1607–1608 by the Jesuit Diego de Gonçález Holguín (1552–1618), (Calvo Pérez 2004). A helpful survey of these and other early colonial materials on Quechua is in Mannheim (1991:138–152). The most prolific writer on Aymara, another widespread Andean language, was the Italian Jesuit Ludovico Bertonio (1552–1625), who wrote three grammars and a dictionary, as well as several religious works. Bertonio claimed to follow the Latin model for his morphology, but to have departed from it in his syntax (Briggs 1985:548–549, Rowe 1974:365). However, it is also the case that original descriptive models and terminology arose in Spanish America, and spread from one region to another, as demonstrated by Adelaar (1997).

Laughlin and Haviland (1988:8–27) surveyed 16 dictionaries of Latin American indigenous languages compiled in the sixteenth, seventeenth and eighteenth centuries by Spanish missionaries. They note a tradition of modeling dictionaries of one indigenous language upon another, but not necessarily upon a Spanish model, and report that the dictionaries of Mexican languages were more thorough than those of Quechua or Guaraní of South America.

In North America, Guale of Florida was the first language to be written down by the Spaniards. We know that the Jesuits worked on it (1668–1670), and then the Franciscans (1678–1763), but the documentation has not survived (Sturtevant 2005:11). The best documented language by the Spaniards in Florida was Timucua. Francisco Pareja, who died 1628, wrote an extensive Timucua *arte* and a *confesionario*, described in some detail by Sturtevant (2005:11). Both Guale and Timucua are now extinct.

In Brazil, the first missionary fieldworker was the Jesuit José de Anchieta (1533–1597), who described the Tupinambá language of the coastal areas around São Paulo and Rio de Janeiro (Rodrigues 1997). The French Calvinist Jean de Léry (1534–1613) is notable for having recorded natural Tupinambá conversations, the only record we will ever have of conversations in this language (Rodrigues 1997:376–383, Everett 2004). South and south-west of Brazil, missionary activity was also very productive. The Franciscan Luís de Bolaños wrote the first Guaraní grammar, vocabulary, and prayer book, and portions of catechism (between 1583 and 1585), but all of these were lost.[3] The Franciscan Francisco Solano worked on Chaco languages (Caraman 1976:26–27).

[3] As we will see in this account, it is not unusual for important unpublished fieldwork materials, to get lost, for a variety of reasons. Hymes (1963:71) provides several more cases from the twentieth century.

The Franciscans were then eclipsed by the Jesuits, who founded a number of republics (1609–1767) in what is now eastern Paraguay and adjacent regions of northern Argentina and southern Brazil (Ostler 2005:371). These Jesuits were impressive field linguists (Landar 1976b). Prominent names are the following: Antonio Ruiz de Montoya (1585–1652), Procurator[4] of the Paraguay province in 1636, wrote a superb grammar, a dictionary, and other works on Guaraní (Grannier Rodriguez 1997), and was instrumental in the propagation of a Guaraní culture and language in modern-day Paraguay (Caraman 1976:294). Martin Dobrizhoffer (1717–1791) wrote on the Abipón language. José Brigniel wrote an Abipón dictionary ca. 1650 (Caraman 1976:198, 206, 209). Sánchez Labrador wrote grammars and dictionaries of Guaycurú and Mbayá, (ca. 1760, published only in 1916), and was also an excellent naturalist and ethnologist (Caraman 1976:202–203). Antonio Machoni de Cerdeña, the Jesuit Provincial of Paraguay[5] from 1739 to 1743 (Caraman 1976:311, 314), wrote an *arte*, a *vocabulario* and a *doctrina* of the extinct Lule, the only existing documentation for this language (Machoni 1877), originally published in 1732.

Upon their expulsion from Paraguay in 1767, the Jesuits were not allowed to take their papers and books with them, and many fieldnotes were undoubtedly lost. For example, we know that Ignacio Chomé wrote on the Chiquito language, but his manuscripts have been lost (Caraman 1976:277, 281).

The writing of grammars and dictionaries in the Spanish missionary tradition slowed down in the seventeenth century, but still continued until the mid-eighteenth century (McQuown 1976:106–107). Accounts of Spanish missionary language work in the sixteenth and seventeenth centuries are in Campbell (1997:30–31) and in Zwartjes (2000).

3.2.3 Missionary Beginnings in North America

In North America, the first language documentation dates from Jacques Cartier, regarding Micmac (Algonquian) in the Quebec area (1534–1536). The activities of missionaries in New France on Algonquian and Iroquoian languages in the seventeenth and eighteenth centuries are aptly described by Hanzeli (1969). Hanzeli (1969:32–44) explains how French missionary grammatical work was due to the Latinate grammatical training of Catholic (and especially Jesuit) missionaries. From the point of view of the historiography of fieldwork, Hanzeli's (1969:45–54) discussion of missionary fieldwork and language learning is particularly revealing. It shows how successive generations of missionaries tended to build on and improve upon the unpublished materials of their predecessors, and that some missionaries

[4] According to Caraman (1976:312) a Procurator is "A delegate chosen by the Provincial Congregation to conduct the affairs of the Province in Europe or to represent its interests at a General Congregation."

[5] In Jesuit administrative terminology, a Provincial was a priest appointed to take charge of a Province or Jesuit Administrative unit for a period of 3–6 years.

were much better language learners than others. Also, in contrast to what happened in Spanish America, where many grammars were published at an early date, the large majority of French missionary fieldwork of the period remains unpublished (Rowe 1974:369, Ostler 2004:39–40, Koerner 2004:59).

For the colonial period in English-speaking North America, two missionary New Englanders stand out. Roger Williams (1603–1683), learned the Narragansett language by immersing himself among the native communities, and wrote a phrasebook of the language (Williams 1643), the first phrasebook of any North American indigenous language, and full of interesting ethnographic data. The Puritan John Eliot (1604–1690) learned the Massachusett or Natick language of Massachusetts, and wrote a grammar based on consultant work, *The Indian Grammar Begun* (Eliot 1666). This is the first published attempt to describe the language in its own terms (Miner 1974:170, Hoijer 1976:3). Koerner (2004) is an outline of missionary linguistics in North America.

3.2.4 German Colonial Missionary Work in Africa

Because of the philological tradition present in nineteenth century Germany, German explorers and missionaries wrote many contributions to African linguistics, even in those African countries not under German colonial administration. The German missionary Johann G. Christaller (1827–1895), the founder of West-African linguistics (Jungraithmayr and Möhlig 1983:62) understood the tonal phenomenon of downstep (Christaller 1875), which was redescribed by American linguists in the mid 1960s (Pike 1975:11–12).

In Freetown (present-day Sierra Leone), German missionaries were employed by the Church Missionary Society (C.M.S.), which was founded and supported by the Church of England: Jacob Friedrich Schön (1803–1889) (who anglicized his name as James Frederick) was the founder of Hausa studies (Hair 1967:37–41). His successor was Sigismund Wilhelm Koelle (1823–1902), founder of Kanuri studies. Koelle admitted that his Kanuri studies, while of linguistic interest, were not then of missionary use, since Islamic powers would not allow Christian missionaries in the Kanuri homeland itself. This might be one reason why Koelle was transferred to the Middle East and wrote on Turkish (Hair 1967:41–44). This is an example, not unusual, of a missionary fieldworker carried away by linguistic rather than evangelistic pursuits.

These two early scholars had a rather dim understanding of the importance of tone in African languages such as Hausa and Kanuri (Hair 1967:55–56). This is a bit surprising since Koelle participated in discussions on the marking of tone in Romanized Chinese. Tones in Chinese had been recognized by European students as far back as the sixteenth century, most likely because the Chinese themselves had written on tone. It was Bishop Samuel Crowther (ca. 1806–1881), a native Yoruba scholar, who insisted that tones were an important element of the Yoruba language and therefore must be marked (Hair 1967:56–57, 91–92).

3.2.5 *Missionary Work in Australia, Papua New Guinea, and the Pacific*

The most proficient early missionary fieldworker in Australia appears to have been Lancelot Edward Threlkeld (1788–1859), who wrote on the Awabagal language, spoken north of Sydney (Carey 2004). However, except for a few missionaries like Threlkeld, there was hardly any fieldwork on Australian languages between the mid 1840s until the late 1870s (Dixon 1980:12, McGregor 2008a:10–11).

Probably with the exception of Amazonian languages, the languages of Papua New Guinea were the last group of languages to be studied by missionary fieldworkers. In the colonial period, some work was carried out by mostly German or Dutch missionaries, depending on the area of Papua New Guinea (Foley 1986: 12–14). Mühlhäusler (1999) stresses the importance of looking at missionary archives for creole and pidgin studies (particularly for Melanesia) since missionaries tended to be observant of any form of speech.

3.2.6 *The Interesting Case of Moravian Missionary Work*

Moravian missionaries were particularly proficient linguists. Johann Jacob Schmick (1714–1778), was an east Prussian who served on the Pennsylvania missions to the Mahicans, an Algonquian group originally from eastern New York state, and immortalized through James Fenimore Cooper's novel *The Last of the Mohicans* (1826). Schmick wrote a German-to-Mahican manuscript dictionary, reworked by Masthay (1991) into an English-Mahican-German dictionary. While the material in the dictionary is rich in morphological and syntactic detail, Schmick's spelling is not phonemic.

John Gottlieb Ernestus Heckewelder (1743–1823) was born in England of German-speaking parents, and wrote perceptively about the Delaware, an Algonquian group of Pennsylvania (Campbell 1997:35, 381).

Samuel Kleinschmidt (1768–1886) wrote a grammatically and phonologically sophisticated grammar of Greenlandic Eskimo (1851), not based on Latin-based models, but revolutionary in its own terms, more so than John Eliot (see Section 3.2.3). Kleinschmidt was born and died in Greenland; his father was a German and his mother a Dane, and he learned Greenlandic growing up with native-speaking children (Rosing 1951:63). However, since Kleinschmidt was based in Greenland, his grammar had little impact on linguistic theory.

Another Moravian missionary, Johann Heinrich Theodor Bourquin (1833–1914) worked on the Inuit language of Labrador, Canada. His son Walther (1879–1974) became a missionary in South Africa, and wrote on Bantu and Khoisan languages (Jungraithmayr and Möhlig 1983:55–56).

A Moravian missionary to British Lahoul, Heinrich August Jäschke (1817–1883) wrote the first scientific dictionary of Tibetan (1881) (van Driem 2001:848).

With the Tibetan orthography and grammatical tradition to depart from, Jäschke's fieldwork was not as groundbreaking as Kleinschmidt's. His genius is that, in view of the goal of translating the Bible into the colloquial language, he made an effort to include information on the various spoken dialects, hitherto ignored by the native tradition. Another Moravian Tibetologist was August Hermann Francke (1870–1930), an authority on Ladakhi dialects (van Driem 2001:935).

3.2.7 Other Nineteenth Century and Early Twentieth Century Missionary Work

In British India, the East India Company long forbade proselytizing, and Christian missions were only allowed to work with a license beginning in 1813, and freely in 1844 (van Driem 2001:470). American Baptist missionaries predominated in northeast India, and by the beginning of the twentieth century had carried out quite a bit of fieldwork and Bible translation in the Tibeto-Burman languages of what are now Arunachal Pradesh, Manipur, Meghalaya, Mizoram, and Nagaland (van Driem 2001:469, 471, 529, 574, Coupe 2007:14–15). American Baptist missionaries also wrote on the language of the Chin Hills in Burma (van Driem 2001:586).

There were several Dutch fieldworkers in what is now Indonesia. The most skillful was Hermanus Neubronner van der Tuuk (1824–1894), a missionary with the Dutch Bible Society and the first to study the Toba-Batak language of Sumatra. His monumental grammar (van der Tuuk 1864, 1867) is still the best description of the language. Percival (1981:3–5) points out that the syntax is not systematically presented, and that van der Tuuk apparently did not elicit spoken Toba-Batak, but had his informants write down texts for him, which he then analyzed. Nevertheless, one particularly perceptive quotation from the English translation (van der Tuuk 1971:xliii) well compensates for these shortcomings:

> I do not believe that anyone will ever be able to represent a language well if he does not disabuse himself of the striving for a complete system, for every language is more or less a ruin, in which the plan of the architect cannot be discovered, until one has learnt to supply from other works by the same hand which is missing in order to grasp the original design.

The British missionary Thomas Bridges (ca. 1842–1898) was the first permanent non-indigenous resident of Tierra del Fuego. He wrote a Yamana (Yahgan) dictionary (Bridges 1933) which is quite remarkable for the detail and precision of the English definitions. One of Bridges' sons wrote a very interesting autobiography and ethnographic account (Bridges 1951), pointing out that much of Thomas Bridges' fieldwork on Yahgan was carried out on the Falkland Islands, with speakers who had been left behind there by a ship (Bridges 1951:46–48). This book also shows that Bridges' sons were quite proficient not only in Yahgan but also in Ona or Shelknam, another unrelated Fuegian language.

In North America, some of the best fieldworkers of the nineteenth century were Catholic and Protestant missionaries (Hoijer 1976:7, Sturtevant 2005:35–42). The Oblate Adrien Gabriel Morice (1859–1938) wrote a monumental grammar and dictionary of the Carrier language, but he was harsh on practitioners of native religions and did not get along with other priests either. Morice also had strong opinions about native contributions to phonetic transcription; he believed that indigenous intelligence was not flexible enough to be of help in such matters (Morice 1904:240–241). The Jesuit Francis Barnum (1849–1921), wrote an ethnographically sophisticated grammar, as well as an account of fieldwork among the Yupik Eskimo of Southwestern Alaska (Barnum 1893).

The Jesuit Jules Jetté (1864–1927) probably did not have as good an ear as Morice, but he wrote a monumental and ethnographically very rich dictionary of the Koyukon language of western Alaska, now published as Jetté and Jones (2000).

Farther south, many Catholic missionaries wrote on the languages of Latin America. A few examples are the Dominican José Pío Aza (1867–1938), who wrote on languages of the Peruvian Amazon (Junquera Rubio 2005), the Scheutist[6] Esteban Haeserijn V. (1913–1975) who wrote on Q'eqchi' of Guatemala (Haeserijn 1966), and the Capuchin Camilo Múgica who wrote on Guajira of Venezuela (Múgica 1969).

3.2.8 The Summer Institute of Linguistics and the Wycliffe Bible Translators

The U.S.-based Summer Institute of Linguistics (now called SIL International), the missionary arm of which is known as the Wycliffe Bible Translators, was founded by the evangelical missionary William Cameron Townsend (1896–1982). Townsend's biography is Hefley and Hefley (1984). A somewhat romanticized account of SIL activities, which is nevertheless worth reading, is Wallis and Bennet (1964). Harsh critiques of SIL activities in Latin America are given in Kummer (1981:179–183) and Stoll (1982). A more balanced discussion of the influence -political and linguistic- of SIL is in Newmeyer (1986:59–61), who views it is a powerful American structuralist force. A balanced discussion of SIL activities in Nepal is in van Driem (2001:790–794). Bright (1967) is a good survey of materials for Mesoamerican descriptive scholarship until 1966, which gives a good feel for the work done by SIL as compared to work by others.

---

[6]The Scheutists, also called C.I.C.M. (Congregatio Immaculati Cordis Mariae "Congregation of the Immaculate Heart of Mary") are a Belgian Roman Catholic missionary organization.

Superb fieldworkers with connections to SIL or Bible translation are Henry Allan Gleason Jr. (1917–2007), Eugene Nida (born 1914), William Smalley (1923–1997), and Kenneth L. Pike (1912–2000). Pike, who was influenced by Sapir, was certainly the best fieldworker affiliated with SIL, and possibly the best field linguist in the history of the field. He is responsible for many practical methodologies and analytical techniques for fieldwork (Pike 1947, 1948). He was so proficient in these that he often gave demonstrations, working monolingually with a speaker he had never met, and was able to give a believable sketch of the phonology of a language after just 45 min. He was doing such demonstrations in lectures at least as early in 1945 (Hymes and Fought 1981:119). Makkai (1986) contains a description of such a demonstration, which Pike continued to do until shortly before his death.

SIL's role in field linguistics is controversial, as can be seen in the recent debate on the role of SIL within the disciplinary culture of linguistics (Dobrin 2009; Dobrin and Good 2009; Svelmoe 2009; Handman 2009; Epps and Ladley 2009; Olson 2009). The debate is not regarding whether SIL linguists are poor fieldworkers, as they clearly are not, but about whether religious proselytization is inherently destructive of indigenous cultures and something that could be considered an unethical activity (see also Chapter 6). The extent to which proselytization and conversion destroys indigenous culture differs around the world, and blanket condemnation of SIL activities seems based on ideology and is oftentimes unconnected to actual activities in specific field sites.

3.2.9 Lessons from This Section

What we can learn from this is that there always have been extremely talented and perceptive missionaries in the field, regardless of what we might otherwise think of their activities. Because of their intimate and long-term involvement with the native people, they often learned the language well; but that of course does not mean that they were good fieldworkers, or that they wrote down interesting things about the language. Literature by certain orders of missionaries such as the Moravians and Jesuits turns out to be, in general, somewhat more perceptive than that of other groups, but many other groups, like the Capuchins (O.F.M. Cap.), Dominicans (O.P.), Franciscans (O.F.M), Oblates (O.M.I.), Scheutists (C.I.C.M.), Lutherans, and various other Protestant missionary societies based in the Netherlands, England, Scotland, or Germany, also did valuable fieldwork. It is not surprising, then, that there has been a recent resurgence of interest in missionary linguistics, as seen in the useful collections by Zimmerman (1997a); Zwartjes and Hovdhaugen (2004), in particular the contributions by Zimmermann (2004) and Ostler (2004); Zwartjes and Altman (2005); Zwartjes et al. (2007, 2009). A useful survey article on missionary linguistics from 1500 to 1900 is Gray (2000).

3.3 "Gentleman Scholars" and Linguistic Fieldwork

Overlapping with missionary fieldwork to some extent are the activities of people we will call "gentleman scholars".[7] Either gifted amateurs or scientists in other fields, they were explorers or travelers, military or colonial personnel, ambassadors, politicians, or simple farmers or school teachers. These people were not doing fieldwork as a primary activity, but did it out of interest or humanist curiosity, or as a hobby, and earned their living through other activities.

In Africa, the study of indigenous languages started with the first wordlist, compiled in 1479 or 1480 by the Flemish traveler Eustache de la Fosse what is now Guinea. It consists of seven words, apparently from Twi (Gregersen 1977:92).

The first Australian Aboriginal word to be written down by a European is found in the writings of the privateer William Dampier, who heard it in 1688 (McGregor 2008b:2). The first attempt at a systematic documentation of an Australian Aboriginal language were word lists of Guugu Yimidhirr by members of Captain Cook's expedition in 1770 (Dixon 1980:8). Similarly, the first documentation of a Papuan language was a word list of Kamoro written in 1828 by a passenger on a Dutch ship (Foley 1986:12–14).

More intriguing and revealing cases of "gentleman scholar" collecting are treated in Sections 3.3.1–3.3.6.

3.3.1 Busbecq and Crimean Gothic

An intriguing case of early fieldwork is that of the Flemish nobleman Ogier Ghiselin de Busbecq (ca. 1520–1592) who served Ferdinand I of Austria as Imperial Ambassador to Suleyman the Magnificent. In his Turkish Letters, written in 1562 and first published in 1586 (Busbecq 1968), he provides a list of "German" (i.e. Germanic) words, including four phrases and 18 cardinal numbers, and three lines of a song without translation - in all 101 separate forms. Busbecq had heard of people in the Crimea who spoke a "German" language, and when he had an opportunity to meet envoys from that area, he sent his interpreters to bring them to his residence for dinner. Of the two envoys, one looked Flemish or Dutch, but had completely forgotten his language. The other was Greek by birth, and a native speaker of Greek, but from dealing with the Germanic speakers "had acquired a fair acquaintance with their language" (from Stearns' translation from the Latin). So the Greek was questioned by Busbecq. Busbecq could not make up his mind whether these Germanic people were Saxons or Goths.

The German words recorded by Busbecq have long been identified as samples of Gothic, indeed the only Crimean Gothic words ever recorded, and the very last

[7] The term "gentleman scholars" accurately reflects a time when women were generally excluded from linguistic fieldwork pursuits.

words ever recorded of an East Germanic language. Busbecq apparently only wrote down the words that reminded him of his native Dutch (Stearns 1978:9–15), and it is unfortunate that he did not write down all the words his interlocutors provided him with. On the basis of reports (without language data) following Busbecq’s report, Stearns (1978:15–19) assumes that Crimean Gothic survived till the last decade of the eighteenth century. If it had not been for Busbecq’s curiosity, we would have no data on Crimean Gothic whatsoever.

3.3.2 Hennig von Jessen and Polabian

Polabian, a Western Slavic language that was spoken in several villages west of the Elbe river in northern Germany, became extinct somewhere in the middle of the eighteenth century. The best records were compiled by Christian Hennig von Jessen (1649–1719), a local pastor, who in the introduction to his vocabulary (published in a facsimile edition by Olesch 1959) explains the difficulties he had in obtaining consultants: the language was almost extinct, people who knew some of it were ridiculed, and no one wanted to admit that they knew it. Polabians, being a peasant people, were always working and were not interested in words, and as a preacher Hennig was busy with services on Sundays, their only free day. He ultimately secured the help of a peasant. Schleicher used Hennig’s work as well as a few other compilations to write a grammar (1871), one of the first philological interpretations of fieldnotes of a moribund language. Hennig had a good ear, and used accent marks (Polański and Sehnert 1967:15–19). However, his account does contain some of the very typical errors which later interpreters of fieldnotes have to contend with; when he asked for ‘yesterday’, ‘today’, and ‘tomorrow’, the speaker answered with ‘Friday’, ‘Saturday’, and ‘Sunday’, which was unquestioningly written down by Hennig (Pedersen 1931/1962:51–52, Polański and Sehnert 1967:193).

3.3.3 Lhuyd, Barrington, and the Last Speakers of Late Cornish

The Welsh scholar Edward Lhuyd (1660–1709) wrote on the living Celtic languages, and spent 4 months documenting the moribund Cornish (Lhuyd 1707). He had the good sense to write Cornish using not a spelling system, but a phonetic system. He expressed distaste for some of the (possibly English influenced) features of the language. This disdain for the language appeared to be shared by the last speakers as well (Pedersen 1931/1962:11, Wmffre 1998:3, 5).

The English lawyer, antiquary and naturalist Daines Barrington (1727–1800) found a few native speakers of Cornish at the town of Mousehole, including Dolly Pentreath, who died at the age of 102 in 1777 and who is traditionally considered the last speaker of the language (George 1993:414). She actually was not the last speaker, but was probably the last speaker raised as a monolingual. It is interesting that at her death she underwent a mythologization as the last speaker, but no one seems to have thought of working with her as a consultant while she was still alive. The actual last native speaker died in 1891 (Grenoble and Whaley 2006:46).

3.3.4 Imperial Britain

Imperial Britain produced several extraordinary fieldworker missionaries or colonial administrators. An early grammar that stands out for its attempts at describing the language in its own terms is the work of the Irish Orientalist William Marsden (1754–1836), who worked for the British East India Company. His grammar of Malay (Marsden 1812) was praised by Wilhelm von Humboldt (Hymes 1963:84).

Edgar Horace Man (1846–1929) was a British colonial administrator for the Andamans, and was apparently an obsessive collector of words and objects. His remarkable *Dictionary of South Andamanese* (Man 1923) is English-Andamanese only, but is extremely rich in ethnographic and linguistic detail. This unusual work deserves a fuller analysis. Another devoted Victorian was the missionary Thomas Bridges, discussed in Section 3.2.7.

The colonial administrator and Irish philologist Sir George Abraham Grierson (1851–1941) compiled the 19 volume Linguistic Survey of India (1903–1928). The project began in 1894, and was finished in 1927, and was intended to cover British India; however, for unclear reasons, the whole south of India was excluded (Singh 2006). Local government officials were asked to respond to printed questionnaires and to supply samples both of original texts and of the parable of the Prodigal son (New Testament, Luke 15:11–32), and almost all of them complied. The use of this parable was a common practice of nineteenth century dialectological elicitation in European languages.

The colonial administrator Sir Charles James Lyall (1845–1920) best known as an Arabic scholar, also studied Mikir (or Karbi), a Tibeto-Burman language of Meghalaya and Assam, and contributed Mikir material to Grierson's survey.

The colonial administrator Lt. Colonel D. L. R. Lorimer compiled a grammar, text collection, and dictionary of Burushaski, a language isolate of Pakistan and Kashmir (Lorimer 1935a, b, 1938). According to Emeneau (1940), "The grammar can be said safely to be one of the best "amateur efforts that has ever appeared in linguistics, and is marked by acuteness both of recording and of analysis".

Other noteworthy British military commanders who took an interest in the local languages include Colonel John Davidson, who wrote important work on Kati, a Nuristani language of Afghanistan (van Driem 2001:1084); Colonel George Byres Mainwaring (1825–1893), who provided important early documentation of Lepcha, a Tibeto-Burman language of Sikkim, although he tried to use the categories of Latin to describe Lepcha grammar (van Driem 2001:824–825); and Colonel Leslie Waterfield Shakespear (1860–1933), who wrote the first book of tales in Lushai (or Mizo), a Tibeto-Burman language of Mizoram, India (van Driem 2001:585).

As far as Australia is concerned, Dixon (1980), in his chapter on the history of ideas about Australian languages, notes that the period 1875–1910, was one of relatively benign treatment of Aborigines, and that in this period there was a resurgence of interest in Australian languages, especially by farmers and other amateurs.

3.3.5 German "Gentleman Scholars" in the Nineteenth Century

German traveler-researchers also tended to be thorough fieldworkers.[8] Ernst W. Middendorf (1830–1908), who arrived in Peru as a medical doctor and was later a landowner, wrote on the native languages of Peru, combining philology with fieldwork (Adelaar and Muysken 2004:18–19). Otto Dempwolff (1871–1938) was a doctor and a colonial administrator in German New Guinea and later in Southwest Africa and East Africa. He carried out perceptive fieldwork on Austronesian languages as well as on Khoisan languages. The botanist and poet Adelbert von Chamisso (1781–1838) wrote on the grammar of Hawaiian. Karl Friedrich Philipp von Martius (1794–1865) and Karl von den Steinen (1855–1929) worked on languages of Brazil, and had comparative interests (Campbell 1997: 54, 80). The superlative Brazilian fieldworker Curt Unkel (1883–1945) was of German origin, but took the Guaraní name Nimuendajú, and lived and died like an Amazonian native.

3.3.6 Commissioned Wordlist Gathering

An activity which stimulated fieldwork by "gentleman scholars" was the gathering of wordlists commissioned by powerful patrons, philosophers, politicians, or even royalty. The Empress Catherine II (Catherine the Great, reign 1762–1796) had word lists and specimens collected from the vast Russian empire, much of it by the German biologist Peter Simon Pallas (von Adelung 1976; Ivić 1965:33; Klein and

[8] One should not lose track of the fact that, notwithstanding the opportunities for travel afforded by missions and colonies, not all nineteenth century study of exotic languages implied fieldwork. In addition to Schuchardt (Section 3.6.1.2), two other notable German-speaking linguists interested in exotic languages never carried out fieldwork themselves. The most famous of such armchair linguists was probably Wilhelm von Humboldt (1767–1835), well known for his interest in universals, linguistic relativism and language as reflection of world view (Humboldt 1971). He was the first scholar to study Javanese, an Austronesian language, and did this without conducting any fieldwork whatsoever, thus on the basis of materials collected by others (Ivić 1965:48–50; Percival 1974). Another armchair linguist was Georg von der Gabelentz (1840–1893), who wrote about grammar-writing (von der Gabelentz 1891), and who may have been an influence on Saussure (Coseriu 1967), although the extent of this influence is debated by Koerner (1975:791–792). Von der Gabelentz continues to influence modern German fieldworkers (Mosel 1987, 2006; Zaefferer 2006), and his potential influence on the Boasian tradition would be well-worth exploring.

Klein 1978; Evans 2010:35). Dedenbach-Salazar Sáenz (2006) is a discussion, with useful bibliography, of a Quechua vocabulary collected for Catherine II, written in 1788 by someone was a perceptive fieldworker, and who might have been a native speaker of Quechua.

Other compilations were those by the Jesuits Filippo Salvatore Gilij (1721–1789), and Lorenzo Hervás y Panduro (1735–1809), based on unpublished data provided by their missionary colleagues (Hymes 1963:66; Landar 1976b:186; Suárez 1983:5; Campbell 1997:31–34), as well as the *Mithridates*[9] *oder allgemeine sprachenkunde, mit dem Vater Unser als Sprachprobe in beinahe fünfhundert Sprachen und Mundarten*, in four volumes (1806–1817), by Johann Cristoph Adelung (1732–1806) and Johann Severin Vater (1771–1826) (Jespersen 1921:22; Pedersen 1931/1962: 2, 10; Ivić 1965:38).[10]

More linguistically sophisticated wordlist compilations were *Asia Polyglotta* (1832) by the Prussian Orientalist H. Julius Klaproth (1783–1835) (Pedersen 1931/1962:100), and *Polyglotta Africana* (1854) by the German missionary Sigismond Wilhelm Koelle (1823–1902) (Gregersen 1977:95).

The U.S. president Thomas Jefferson (1743–1826) was interested in Native American languages. He made up a 250 word list and had vocabularies sent to him. He collected 50 vocabularies from actual informants over a 30 year period. Unfortunately, most of his material was lost in 1809 (Sturtevant 2005:19).

Before museum-sponsored fieldwork started in earnest, the Americans Peter S. Duponceau (1760–1844), and Albert Gallatin (1761–1849) had devised elicitation schedules to be sent around by the Secretary of War for gathering information on Native American languages (Hoijer 1976:6; Landar 1976a:93–94). Gallatin and Duponceau's work is further described for Southeastern U.S. languages in Sturtevant (2005:21–29).

3.3.7 Lessons from This Section

What modern fieldworkers can learn from this is the following: if they are dealing with an unusual or endangered language, they should try to collect as much as they can. It will not do to write down only what is of direct interest to the fieldworker. Furthermore, the notes should be archived properly so they will not be lost to posterity (see also Section 5.2 on philology). It is quite clear that the "gentleman scholars" had very little or no phonetic training, and sometimes a very poor ear; the modern fieldworker is expected to do better, but it should be a comfort to know that these amateur transcriptions have had their influence on modern linguistics; an imperfect transcription is better than none at all.

---

[9] The name of King Mithridates of Pontus (first century BCE), who was said to have known twenty-five languages, seems to have been popular as a title of multilingual compilations.

[10] A detailed overview of early word collecting, with particular reference to Finno-Ugric, is in Gulya (1974:258–267).

3.4 Fieldwork in Less Than Optimal Circumstances

Here we list some cases of fieldwork in unpleasant circumstances, either because the consultants were slaves, prisoners, or hospital patients, or because fieldworkers themselves were imprisoned or exiled. Such fieldwork, at least in those instances where it is not unethical, could also be called serendipitous fieldwork.

3.4.1 Slaves

It is a striking fact that the first collection of West African vocabularies (containing Akan, Ewe, Ibo, and Ibibio) ca. 1760, was collected from slaves not in Africa, but in the West Indies (Hair 1967:5).

The British Navy started to suppress the Atlantic Slave trade after 1807 by capturing slave-ships and releasing the slaves at Freetown. It turned out that the majority language of these slaves was Yoruba, but Freetown was a thousand miles away from Yoruba country. As a result, fieldwork on the Yoruba language started with the missionaries based in Freetown. Samuel Crowther (already mentioned in Section 3.2.4) was an ex-slave originally from Nigeria; he became the first Yoruba informant to these missionaries (Hair 1967:6–8, 17).

Of the 17 early vocabularies of Hausa collected in the period 1840–1850, two were collected from African slaves in Brazil, two from ex-slaves in Sierra Leone, one from slaves at Algiers, and one from slaves in the West Indies or on Fernando Po Island; the others were collected by explorers or from merchants (Hair 1967:34–36).

3.4.2 Prisoners

There are two cases of Apaches interrogated during U.S. military campaigns against them in the nineteenth century. Captain John Gregory Bourke (1846–1896), for all practical purposes the scientist on George Crook's campaigns against the Chiricahua Apache, collected vocabularies and grammatical notes on Western Apache, from the Apache scouts who were hired by Crook.[11] Bourke also managed to collect data from the Chiricahua prisoners of war, interestingly mostly women; it appears that the male Chiricahua prisoners of war were in no mood to provide

[11] Many of the words collected from such scouts were Spanish, not Apache, and were provided by the Apaches probably due to their unwillingness to share genuine Apache equivalents. Bourke realized this but wrote the Spanish words up anyway, perhaps realizing that this would some day be an interesting record of the Spanish knowledge of Apache scouts (see Section 5.2).

linguistic information (Bourke 1980). A full account of Bourke's colorful military and scientific career is in Porter (1986).

The celebrated Chiricahua Apache chief Mangas Coloradas met with John R. Bartlett, United States Boundary commissioner in 1851 and Bartlett collected a brief vocabulary from him (Sweeney 1998:232). This is the earliest record of the Chiricahua language, and the only one from a famous Apache leader (Bartlett 1851). Mangas Coloradas was later made a prisoner by the U.S. military, and was murdered by his captors in 1863 (Sweeney 1998:457).

Good fieldwork-based grammars were written by German Caucasianists such as Baron Peter Karlovič Uslar (1816–1875), Anton Schiefner (1817–1879), and Adolf Dirr (1876–1930) (Schulze 2005:322–323). The German and Austrian tradition of Caucasian linguistics was influenced by the presence of prisoner-of–war camps in World Wars I and II. A supply of Tsarist or Soviet prisoners of war were available as informants. For example, in World War I, Adolf Dirr and Robert Blechsteiner (1891–1954), an Austrian Caucasianist, were involved, as Schulze (2005:323) puts it, "in this type of dubious work".

The German scholar and explorer Wilhelm H. I. Bleek (1827–1875), when unable to carry out fieldwork in isolated areas of South Africa for health reasons, studied the language of San (Bushmen) prison inmates (Jungraithmayr and Möhlig 1983:54).

The most unethical practice by field workers in a colonial context is probably corporal punishment. Amazingly, the famous anthropologist Bronisław Malinowski (see Section 3.4.4) admits in his dairies to have resorted to physical violence against recalcitrant consultants (Malinowski 1989 [1967], mentioned in McGregor (2008c:426). I (de Reuse) have heard from a reliable source about beatings of Congolese speakers in the Belgian Congo who refused to provide the correct linguistic information.

3.4.3 Indigenous People in Exhibits, Museums or Hospitals

Traveling nineteenth-century museum exhibits sometimes displayed live human beings, who were treated as curiosities, if not as savage animals, and subjected to examination, measurement, and photography. Poignant (2004:125–126) describes a display of Aboriginal Australians, and the attempts of two anthropologists to study some of their language. It is perhaps not surprising that the Aboriginal people showed some enthusiasm for this sort of activity, since it was a relatively dignified one, compared to other dehumanizing treatments they were subjected to. However, before the anthropologists could get very far in their study, the Australians were shipped off to an exhibit in another country.

For a more recent account of a "captured" - albeit not imprisoned or mistreated - person, we should also mention the sad story of Ishi (ca. 1861–1916) (Kroeber 1961, 1964). Ishi was not only the last speaker of the Yahi language of the Yana group of Northern California, but was also the last survivor of his people. He was discovered (or rather, let himself be captured) in 1911, and worked until his death, at the University of California Museum of Anthropology, which at that time was in San Francisco. He provided information to several scholars, including the linguists Edward Sapir and Alfred L. Kroeber. Kroeber writes:

> Ishi's English was limited for linguistic work, and it was only only Sapir's genius --or, as he put it to me, his "brute memory of corresponding Northern and Central Yana forms"-- that enabled him to salvage Ishi's Yahi dialect. He said it was the most difficult work with an informant he ever did in his life. (Sapir and Swadesh 1960:v)

Hospitals are also a place where speakers from distant areas can be gathered. Michael Krauss (p.c.), intending to make a preliminary survey of Athabascan languages of Alaska in the 1960s, was able to do much of this in a hospital, without having to visit people in isolated areas of Alaska. Another example of the result of fieldwork done with hospital patients from Northern Canada is described in Haas (1968).

3.4.4 Exiled or Imprisoned Fieldworkers

There are a number of cases in which confinement or exile actually provided conducive circumstances for carrying out fieldwork.

The Russians Waldemar Bogoras (Vladimir Germanovich Bogoraz) (1865–1936) and Waldemar Jochelson (1852–1937) were exiled to Siberia for revolutionary activities in Tsarist Russia. This exile did not mean confinement, as these scholars traveled widely and gained a knowledge of Siberian languages. Later they became the experts in expeditions funded by the American Museum of Natural History in New York, and collaborated with Boas. Other exiles of Tsarist Russia were the Ukrainian Lev Y. Sternberg (1861–1927), who carried out fieldwork on Nivkh (Gilyak), and the Pole Bronisław O. Piłsudski (1866–1918), who collected materials on Ainu, both on Sakhalin Island (Bobrick 1992:306–307, Pedersen 1931/1962).

The British anthropologist Bronisław Malinowski (1884–1942) was exiled by the Australian government to the Trobriand Islands for 4 years during World War I (Mead and Bunzel 1960:318). His experiences on the Trobriand Islands may have led him to recognize the importance of participant observation, including his insistence on learning the language as part of the fieldwork process (Hymes 1970:253).

During World War I, Gerhard Deeters (1892–1961), the prominent German Caucasianologist of his time, was himself in a prison camp, where he did fieldwork on Georgian, working with fellow prisoners from Georgia (Schulze 2005:323–324).

During World War II, the French linguist André Martinet (1908–1999) carried out the field research for his first phonological study of varieties of French (Martinet

1945/1971), while a POW in a German prison camp surrounded by other bored French officers from all over France.[12]

3.4.5 Lessons from This Section

What we can learn from this section is the need to respect the modern protections afforded to inmates of prisons or people in hospitals, notwithstanding temptations to circumvent them. We can also see that sometimes a linguist can find a bad situation serendipitous for fieldwork, always remembering, of course, to respect the rights of his or her consultants.

3.5 Fieldwork in Collaboration with Native Consultants

In some cases, native speakers actively collaborated with linguistic fieldworkers, were involved in documenting their own languages, and became literate in their own languages. In most early sources, very little is said about the degree to which the speakers themselves collaborated in fieldwork or became writers of their own languages. Even in the Boas-Sapir-Bloomfield tradition, the names of the native speakers are not always mentioned, and the degree to which they helped shape the record is not always clear.

3.5.1 Native Speaker-Missionary Collaborations

The first native speakers who collaborated with outside fieldworkers received religious educations.

The missionary activities in New Spain resulted in the training of native writers. The most celebrated example is that of Fray Bernardino de Sahagún (1499–1590), who in 1575–1577 encouraged his Nahuatl speaking collaborators to write down their own native oral traditions in their native languages, and was in a sense the "first ethnographer" (McQuown 1967:3). As can be imagined, much of the native religion was destroyed by the Spanish church authorities, and Sahagún's work, which we would now call "salvage ethnography" was frowned upon. We are fortunate that these materials, of inestimable value to our understanding of sixteenth

[12] I (de Reuse) collected data on Flemish dialects and regional Belgian French during my military service (1977–1978), so the melting pot of soldiers from all over the country remains an opportunity for rapid and informal fieldwork.

century Nahua culture and language, have survived (McQuown 1976:113–114, Evans 2010:32–34).

We already mentioned Samuel Crowther (Sections 3.2.4 and 3.4.1), a Yoruba speaker originally from Nigeria, who wrote himself in and about the language, and was the founder of Yoruba written literature. His Bible translation was more idiomatic than those written by foreigners. He also became the first African bishop (Hair 1967:6–8, 17).

Pablo Tac (1822–1841) was a Luiseño speaker from southern California, who went to Rome in 1834 to receive a Catholic education, and wrote a grammar of his native language there (Goddard 1996:42).

Other early Native American writers on their own languages were the Abnaki speaker Joseph Laurent from Quebec, who wrote a remarkable and phonemically accurate phrase book in his language (Laurent 1884), and the Quiché speaker Patricio Xec Cuc (1905-??) from Guatemala, who collaborated with Bible translators and edited a version of the Popol Vuh (Tedlock 1983:131).

3.5.2 Early Cases of Recognition of Native Speaker Talent and Insight

The German philologist August Schleicher (1821–1868) recognized the linguistic abilities of native speakers, by noting that the Lithuanian priest Friedrich Kurschat (Lithuanian Fridrichas Kuršaitis) was the first to study and figure out the rules for the position of the Lithuanian accent (Pedersen 1931/1962:65–66). However, Schleicher had difficulty in distinguishing two of Kurschat's accents, and as a result he was, as Pedersen (1931/1962:66) puts it, "inconsiderate enough to say that Kurschat had probably been too subtle." It turned out that Kurschat's view was completely correct.

The phonetician and fieldworker Daniel Jones (1881–1967) was also one of the first fieldworkers to acknowledge the help and insight of native speakers, as seen in his work on the phonetics of Sechuana (Tswana) with speaker Solomon Plaatje (1876–1932). Plaatje became one of the foremost Black South African politicians and writers of his day, and was a founding member of the South African Native National Congress, the organization renamed as the African National Congress in 1926 (Collins and Mees 1999:160–161).

3.5.3 Native Speaker Linguists in the Boas-Sapir-Bloomfield Tradition

Several native speakers of Native American languages carried out fieldwork on their own languages, and/or collaborated with non-native fieldworkers, working in the Boas-Sapir-Bloomfield tradition (Section 3.1.3). We provide here a partial list of these remarkable scholars, followed by their dates, native language, and names of the linguist(s) they worked with, where known:

1. Ella C. Deloria (1889–1971) (Yankton Dakota) with Franz Boas (see Medicine 1999)
2. Juan Dolores (1880–1948) (Tohono O'odham) with J. Alden Mason and Alfred L. Kroeber (see Mathiot 1991)
3. Edward Dozier (1916–1971) (Tewa) with Harry Hoijer
4. William Jones (1871–1909) (Fox), who was a student of Boas, with Truman Michelson from the Bureau of American Ethnology
5. Francis La Flesche (1857–1932) (Ponca) with the Bureau of American Ethnology
6. William Morgan (1917–2001) (Navajo) with Robert Young
7. Doña Luz Jiménez (ca. 1895–1965) (Nahuatl) with Benjamin Lee Whorf, Robert Barlow, and Fernando Horcasitas (see Karttunen 1991, 2000)
8. Gilbert Natchez (Paiute) with Alfred L. Kroeber
9. Alex Thomas (1895–1971) (Nootka) with Edward Sapir and Morris Swadesh
10. Albert Yava (1888–1980) (Hopi and Tewa) with Edward Kennard

3.5.4 The Impact of Kenneth Hale

Kenneth L. Hale (1934–2001) taught at the Massachusetts Institute of Technology and was a colleague of Noam Chomsky (Section 3.6.2), but unlike Chomsky he was a dedicated fieldworker, who worked on Southwestern U.S. Native American languages such as Jemez, Hopi, Navajo, and Tohono O'odham, and Australian languages such as Warlpiri and Lardil. He was also an unusually gifted polyglot. In keeping with Chomsky's emphasis on introspection, he suggested that native people themselves should be trained as linguists working on their own languages, and that linguistic work on exotic languages would greatly benefit from this situation (Hale 1965, 1972, 1976).

Hale was not only thinking about science but also had a humanistic approach. He saw the collaboration of native-speaking linguists with non-native linguists as a corrective to the traditional situation where the non-native fieldworker would have power over the situation, and gain credit for the findings, whereas the consultant would be powerless and gain no credit. Several Native American linguistics Ph.D.s graduated from the Massachusetts Institute of Technology under the impetus of Hale.

The recent influence of native speaker linguists on Mayan linguistics in Guatemala is a very welcome and successful development, entirely in line with Hale's vision (England 2007). Further discussion of native speaker-linguist collaborations is in Hale (1972:388–395), McQuown (1976:114), Davis (1977), and Mithun (1996:56–58).

3.5.5 Lessons from This Section

What one can learn is that there should have been more acknowledgment, in the course of history, of the contributions of native speaker consultants their collaboration with non-native-speaking fieldworker. One also wishes that more speakers of little studied languages would become linguists, and that more modern fieldworkers would encourage speakers to take this path. However, it remains a fact that most indigenous people who wish to work towards advanced degrees find that they might serve their communities better by becoming policymakers, lawyers, or medical doctors, rather than linguists!

3.6 Fieldwork Contemporaneous with Academic Traditions Less Supportive of Fieldwork

What do we mean by "academic traditions less supportive of fieldwork"? Once linguistics was recognized as a scientific field, goals set for linguistics at particular points of time have influenced the amount of fieldwork considered necessary and the amount carried out.

Thus, in the period oriented towards relativism and empiricism, dominated by Boas, Sapir, Bloomfield and their students – this was roughly the first half of the twentieth century-- intense, high-quality fieldwork was being conducted (see Section 3.1.3). In other periods, such as the time of the neogrammarians (ca. 1870–1920), and the time of Chomsky's dominance in American linguistics (1957 to the present), periods that in some sense were even more definitional of linguistics as a science, the practice of fieldwork was considered less important.

The neogrammarians focused on getting their data from ancient texts, and the Chomskyans primarily get their data from introspection. This does not mean, however, that the techniques of fieldwork were or are not practiced at all during the neogrammarian and Chomskyan periods. During the neogrammarian period it was the dialectologists who did fieldwork, and during the Chomskyan period it is primarily non-Chomskyans, i.e. linguists who continued the Boasian tradition at some universities in the United States, Australia, and Europe, and sociolinguists and ethnographers of speech who continued to do fieldwork.

3.6.1 Fieldwork in Neogrammarian Times

The scientific study of language started with the philological and comparative approaches of the nineteenth century. Since philology entailed the study of ancient texts, and comparative studies required the oldest textual attestations of a language, fieldwork was a low priority, and the concept of the individual speaker as well as the speech community were ignored. Nevertheless, there were a few exceptions to this view of neogrammarians (Hamp 1974:394–395).

3.6.1.1 Comparativist Fieldwork

Fieldwork was carried out on languages that were interesting for comparative reasons but for which there were no texts. This included languages such as Romani (Gypsy) dialects, studied by the Germans August Pott (1802–1887), Franz Nikolaus Finck (1867–1910), and the Slovenian Slavicist F. Miklosich (1813–91) (Pedersen 1931/1962:17, 50). Finck, who is now better known as a typologist, also wrote the earliest fieldwork-based description of an Irish dialect (Finck 1899).

Remarkable fieldwork was carried out on Yeniseian by the Finnish Uralicist M. A. Castrén (1831–1852). The fieldwork on Yakut (Turkic) by the Sanskrit scholar Otto Böhtlink (1815–1904) resulting in *Über die Sprachen der Jakuten* (1851) is, in Eric Hamp's words, "a remarkable specimen of informant-based description for a preliterate exotic language" (Hamp 1974:395). Böhtlink, not surprisingly, was inspired by the descriptive methods of Panini, the Sanskrit grammarian (Pedersen 1931/1962:106–115, Lane 1945:475–476, Hymes 1963:84). Fieldwork on Caucasian and on Yukaghir of Siberia can be said to have started with A. Schiefner (1817–1879) in the mid-nineteenth century, and fieldwork on Basque dialects with Prince Louis-Lucien Bonaparte (1813–1891) (Pedersen 1931/1962: 115, 125, 136).

Prince Louis-Lucien Bonaparte was the third son of Napoleon's brother Lucien Bonaparte. Bonaparte's main interest was Basque, but he also carried out fieldwork on Albanian. Hamp (1974) is a detailed assessment of Bonaparte's Albanian fieldwork. Hamp compares Bonaparte's forms to that of others, and comes to the conclusion that this particular neogrammarian was quite accurate and does not conform to the stereotype of the neogrammarian as a bookworm with no interest in living or unwritten languages.

The German August Schleicher (already mentioned in Section 3.5.2), carried out pioneering fieldwork on Lithuanian, taking down songs and tales in the huts of the poorest of Lithuanian peasants, "under privations and hardships of which the cultivated gentleman of our days has hardly a suspicion". He adds that "the joys of hearing the splendid forms of this language in living use", allowed him to endure such hardships (quoted in Pedersen 1931/1962:65). Schleicher also published specimens of the German dialect of his hometown, and stands out as one of the rare fieldworkers among German philologists (Jespersen 1921:71–72).

There were also three sorts of dissenting voices that acted as correctives to the claim of the comparative philologists that only old texts are of linguistic interest. All were more empirical than the philologists, and to differing degrees emphasized the importance of fieldwork. They are the creolists, the phoneticians and phonologists, and especially the dialectologists.

3.6.1.2 Early Creolist Fieldwork

The founder of creole and pidgin studies, Hugo Schuchardt (1842–1927), explored the possibility of unwritten, mixed languages of low prestige spoken by illiterate non-Europeans. Such varieties can be reliably accessed by fieldwork, but Schuchardt carried out no fieldwork himself. He relied entirely on written documents provided to him by administrators, explorers and missionaries (Schuchardt 1979:viii, Holm 1988:29–30). An overview of the pioneering fieldwork on creoles and pidgins carried out by such people is in Holm (1988:14–36).

3.6.1.3 Early Phonetic and Phonological Fieldwork

The development of phonetics in Britain was certainly related to the need to devise consistent phonemic orthographies for unwritten languages of the British Empire. The phoneticians relied more on precise (articulatory) phonetic description and precise phonetic transcription than did the comparativists. This emphasis was initiated by Henry Sweet (1845–1912), who wrote in the preface to his *Handbook of Phonetics* (1877:v):

> Many instances might be quoted of the way in which important philological facts and laws have been passed over or misrepresented through the observer's want of phonetic training.

Daniel Jones (already mentioned in Section 3.5), probably the model for "Professor Higgins" of Shaw's *Pygmalion* (1916/1941), (although only Henry Sweet is mentioned in Shaw's preface to *Pygmalion*) and the most prominent phonetic fieldworker in the first half of the twentieth century, continued this tradition (Collins and Mees 1999:155–163). Phonetic fieldwork-based studies of African languages in this tradition are Doke (1926) and Westermann and Ward (1933), and the methodologically very cautious and explicit studies by Jones and Plaatje (1916) on Tswana, and Beach (1938) on Hottentot (Khoekhoe) (Gregersen 1977:96, Collins and Mees 1999:154–163). The fieldwork-based phonetic studies by Peter Ladefoged (1925–2006) and work by Ladefoged's students, such as Traill (1985) continue this meticulous tradition.

To some extent, phonologists trained by neogrammarians were also more interested in fieldwork than their teachers had been. The Russian Nikolay S. Trubetzkoy (1890–1938), the founder of the field of phonology (Trubetzkoy 1958). He was trained by German neogrammarians, and also carried out linguistic fieldwork on Caucasian languages. As early as 1907, he became interested in Paleo-Siberian languages and wrote about them on the basis of traveler's notes. On the basis of correspondence with fieldworkers such as Bogoras he wrote about Chukotkan languages, on the basis of notes by Jochelson he wrote on Yukaghir, and on the basis of notes by Sternberg he wrote on Nivkh. Probably more than other Prague school linguists, Trubetzkoy was aware of the value of fieldwork for linguistic theory (Jakobson 1958:273–278).

3.6.1.4 Early Dialectological Fieldwork

Dialectology or linguistic geography was first practiced by Germans. The idea was to provide confirmation of neogrammarian comparativist principles through fieldwork on unwritten dialects. It turned out the dialects were more variable and the sound laws less exceptionless than predicted by these principles.

Georg Wenker (1852–1911) started the first dialect atlas, the *Sprachatlas des Deutschen Reichs,* for which he gathered material from 1876 to 1887. Its methodological weaknesses were numerous, since the questionnaires were answered by mail by phonetically untrained village schoolmasters (Chambers and Trudgill 1980:18–19). However, the coverage and rate of response was unsurpassed, since he got responses from about 45,000 schoolmasters. Bismarck Germany was an authoritarian state, and the schoolmasters found it in their interest to do as they were told. This first atlas was not printed, but written by hand. The atlas was elaborated on and published in print much later by Wrede et al. (1927–1956).

A good overview of dialectological fieldwork and its history is in Francis (1983:48–103). An outline of the field is given in Chambers and Trudgill (1980: 18–23), which should be read along with the survey of the German dialectological tradition in Barbour and Stevenson (1990:55–65).

So, whereas the neogrammarians were interested in dialectology as an empirical confirmation of their methods, some of the first convincing criticisms of those methods came from dialectologists, who pointed out their lack of attention to geographical and social factors in outlining the evolution of dialects (Ivić 1965:63). In fact, dialectologically oriented fieldworkers, now often ignored by theorists, could have made more contributions to linguistic theory through constructive criticism of fieldwork methodologies.

3.6.2 Fieldwork During the Generative Paradigm

We now turn to the transformational generative period in linguistics (Newmeyer 1980). When Noam A. Chomsky (born 1927) came to the forefront of linguistic theory in the late 1950s, he was justifiably critical of the overly mechanistic neo-Bloomfieldian discovery procedures (Chomsky 1957:49–60). Chomsky and his followers emphasized the importance of native speaker grammaticality judgments and native speaker introspection (Chapter 12) as sufficient for the construction of the grammatical model (Newmeyer 1980).

The Chomskyan program saw the biological unity of a language learning device and a universal grammar underlying that device. Since the grammar of all languages was presumed to be the same, the study of exotic unwritten languages was not seen as necessary and it was thought that the goals of understanding universal grammar could just as well be accomplished by studying well-known languages. Another advantage of using well-known languages was that speakers were readily available and easily trained in providing grammaticality judgments or introspective comments on a language.[13]

In addition, linguists nowadays have the impression that before Chomsky, it was generally agreed upon that informant judgments were not valid in fieldwork. This, however, is a misconception. In fact, Sapir, and his students such as Swadesh, Whorf, Newman, and Pike definitely utilized speaker judgments in their fieldwork and saw nothing wrong in doing so (Hymes and Fought 1981:75, 159–163, 187).

While fieldwork came to be seen as a way of describing language that was no longer satisfactory, the fieldwork-based study of exotic languages did continue. One American organization which has continued documenting languages in the Boas-Sapir paradigm through the Chomskyan years is the Alaska Native language Center (ANLC) in Fairbanks.[14] Also, American sociolinguistics, in some sense the American reincarnation of traditional dialectology (Shuy 1990), has always relied on fieldwork.

Sometimes fieldwork-based description has adapted to the Chomskyan paradigm, as seen by Constantino's (1959) *A Generative Grammar of a Dialect of Illocano*, Postal's (1963, published 1979) *Some Syntactic Rules of Mohawk*, Matthew's (1965) *Hidatsa Syntax*, Williamson's (1965) *A grammar of the Kolokuma dialect of Ijo* (1965), Daly's (1966, published 1973) *A generative Syntax of Peñoles Mixtec*, Lindenfeld's (1969, published 1973) *Yaqui Syntax*, Franz's (1971) *Toward a Generative Grammar of Blackfoot*, and Sohn and Bender's (1973) *A Ulithian Grammar*. These grammars tend to show that syntactic elicitation in the Chomskyan model is a slow process, fraught with various dangers. For example, Matthews (1965), although presented as a groundbreaking study, is notoriously difficult to follow and verify. Lindenfeld (1973) is clearly written and still valuable, but has a preponderance of elicited structures, including a passive with agent, an analysis based on a consultant's misunderstanding of elicitation prompts, but which does not, in fact, exist, as shown by Escalante (1990), a native speaker of Yaqui.

In Australia, the tradition of formal fieldwork started later than in the Americas, and it was less affected by Chomskyan ideas about fieldwork. Dixon (1980:12–17) notes that government attitudes became less enlightened in the period from the beginning of the twentieth century up to the 1960s, and in this period very little work was done, with the exception of Arthur Capell. This is, however, an oversimplification, as pointed out by McGregor (2008b:6–8), since it ignores the contributions of several important fieldworkers during that period, in particular the British-trained German linguist Theodor George Henry Strehlow (1908–1978), the son of Lutheran missionaries, who carried out superb work on Arrernte

[13] It is true, as Newmeyer (1980:48) points out, that of the 28 linguistics dissertations written at the Massachusetts Institute of Technology in the 1960s, 17 dealt with languages other than English. However, of these 17 only three dealt with languages without a long-established written tradition: Thomas Bever's on Menominee, James Fidelholtz's on Micmac, and Richard Stanley's on Navajo (Newmeyer 1980:48).

[14] ANLC was founded by an act of the Alaska State legislature in 1972, under the impetus of its first director, Michael Krauss (1973).

(Moore 2008), or the Swedish linguist Nils Holmer (1904–1994) (McGregor and Miestamo 2008).[15]

Modern fieldwork in Australia was started by Arthur Capell (1902–1986), as evidenced by Capell (1945, 1956), but Capell's major interest was Pacific languages (Dixon 1980:16). Since then Australian fieldwork has been vigorously shaped by Capell's student R. M. W. Dixon, (born 1939) who has also written an engaging and candid autobiographical account of his fieldwork in Australia (Dixon 1984). The importance of Dixon's work is also underscored by Blake (2009), in a very balanced review of McGregor (2008a), which mostly covers pre-sixties linguistic research.

Kenneth Hale (already mentioned in Section 3.5.4) carried out superb and extensive fieldwork in Australia, reported on by his wife in Hale (2001), through an interview (Green 2001), and by fellow linguists (Yengoyan 2001, O'Grady 2001). Hale sent several students to Australia, but his influence there has been more in the realm of theory and language activism than in the realm of fieldwork-based description and documentation (Dixon 2004).

Further discussion of the Australianist tradition and its parallels and connections with the Americanist tradition is in Rigsby (1976), and in the superb treatment of the history of Australian linguistic research in McGregor (2008b).

3.6.3 Lessons from This Section

What we can learn from this section is that the fieldworker should keep doing descriptive fieldwork, even when it is relatively unpopular according to the prevailing linguistic theory of the day. Coming generations will appreciate the work.

Also, it is good to remember that when dialects are disappearing under the pressure of a more prestigious language, the research of the dialectologist resembles that of the modern-day documenter of endangered languages. *Mutatis mutandis*, there are similarities between the disappointments and frustrations experienced by Willem Pée, a Flemish dialectologist looking for endangered Flemish dialects in northern France (Pée 1946:VII-XX), those that the fieldworker Bob Dixon experienced, on the other side of the globe, when looking for endangered Australian languages in Queensland (Dixon 1984), and those that the fieldworker Luise Hercus experienced when looking for the last speakers of Australian languages in Victoria and South Australia (Hercus 2008).

We can also gain an appreciation of the fact that fieldworkers past and present sometimes use concepts without being aware of their theoretical importance. For example, before the birth of the field of phonology, a fieldworker would have developed some unconscious concept of the phoneme, without verbalizing it

[15] Holmer was one of those unusual fieldworkers who had experience on several continents; in addition to his work in Australia, he carried out fieldwork on Scottish Gaelic and Irish dialects, on Basque, on North American languages (Siouan, Algonquian, and Iroquoian), Central American languages (Cuna), and South American languages (Choco and Guajira).

as such. We can see this in work as early as that of the German-Russian Caucasianist Baron Peter Karlovič Uslar (1816–1875) who worked on Awar, Chechen, Dargwa, Lak, Lezgi, and Tabasaran (Koerner 1975:808; Schulze 2005:323). Also, the Swiss dialectologist Jost Winteler (1846–1929) had used phonological criteria in a description of an Austrian German dialect in 1876 (Ivić 1965:132–33).

3.7 Concluding Thoughts

3.7.1 The Role of Africanist Reflection in the History of Linguistic Fieldwork

It is interesting to note that it is the Africanist fieldworkers who were the first to be explicit about fieldwork procedures. Consider Beach (1938) on Hottentot (Khoekhoe) phonetics; Gleason (1961:286–311) on phonemic fieldwork on Ewe; Snyman (1970) on elicitations of Bushman (San) syntax; Lumwamu (1973) on techniques applied to Kikongo (Bantu); or the difficulties candidly (and not always charitably) reported on by Evans-Pritchard (1940:12–15) when dealing with Nuer (Nilotic) consultants.

The two earliest book length manuals on how to do linguistic fieldwork were written by Africanists: Samarin (1967) and Bouquiaux and Thomas (original French edition: 1976), (English translation: Bouquiaux et al. 1992).

It is also Africanists who have tended to have the most reservations about the recent trend to emphasize fieldwork for the preservation of endangered languages, as shown in Ladefoged (1992), Mufwene (1998), and Newman (1998).

Africanists also have tended to be more explicit about the history of language studies, as evidenced by surveys such as Armstrong (1964), Hair (1967), Jungraithmayr and Möhlig (1983), and Doneux (2003), by studies on colonial phrasebooks in the Belgian Congo by Fabian (1985, 1986), and by recent studies on colonial representations of languages in South Africa by Gilmour (2004, 2007), and in Africa in general by Irvine (2008).

It is possible that Africanist linguists, who often started working in a more blatantly colonial or postcolonial context than the American, Asian, or Australian contexts, were more attuned to the inherent inequalities between linguists and consultants, and that this has led them earlier to a self-conscious reflection on what it is like to be a fieldworker.[16] Samarin (1967:17) points out that two famous African statesmen, Kwame Nkrumah and Jomo Kenyatta were at one time linguistic consultants, and concludes that this "should prevent field linguists from being too casual in their obligations."

[16] This perception of ours appears to be contradicted by Childs (2007:2–3), who states that there has been a "fuller critical evaluation" of linguistic projects in areas of the world other than Africa. We think Childs is correct regarding specifically ethical reflections regarding language endangerment, but that when one considers fieldwork in all its facets, including techniques and colonial inequalities, Africanists, on the whole, started reflection earlier.

On the other hand, the optimistic legacy of the Boas-Sapir-Bloomfield tradition, which viewed language documentation as unambiguously beneficial to all concerned, started reflecting on these issues several decades later.[17] This more recent attitude towards the Americanist tradition is reflected with sensitivity by Hill (1999).

3.7.2 The Role of Recent Awareness of Word-Wide Language Endangerment on Fieldwork

The most recent trend in the history of fieldwork is the realization that many languages, particularly indigenous minority languages, are endangered, and that a concerted and global effort needs to be made in documenting them (Hale 1992a, b; Krauss 1992). This realization has led - thankfully, one might say - to a resurgence of interest in descriptive linguistic fieldwork, even in generative circles. Documentation of endangered languages as a valid fieldwork goal is discussed, with extensive references, in Section 2.2.2.2.

Useful historical perspectives on past and present trends in the documentation of indigenous languages are in Cyr (1999) for Micmac (Algonquian) of Canada, Parks (1999) for Pawnee (Caddoan), Seguin and Nyce (1999) for Nisga'a (Tsimshian), and McGregor (2008a) for the languages of Australia.

3.8 Lessons from This Chapter

If we take a historical perspective going back to Antiquity, it is no exaggeration to say that the history of mankind's interest in other people's languages has been marked by a relative lack of curiosity and imagination punctuated by a few exceptionally curious and perceptive individuals. Many of these individuals, like great artists, were uninfluential in their own times but are being recognized by linguists now. And some were more perceptive than others: while some early fieldworkers perceived that it was necessary to describe the language as much as possible in its

[17] This observation does not imply that until recently there were no scholars in the Boas-Sapir-Bloomfield tradition interested in reflection on fieldwork techniques. Certainly, American structuralists like Paul Garvin and Charles Voegelin published several articles on techniques (Garvin 1964; Voegelin 1954a, b, 1959a, b, 1960; Voegelin and Harris 1945, 1952; Voegelin and Robinett 1954; Voegelin and Voegelin 1963), but their work in that area has not been influential. The more influential Dell Hymes has written copiously on the history of linguistics, as is obvious from the references to this chapter, but his interests have been more anthropologically oriented rather than focusing on fieldwork technique. Linguists from SIL also published quite a bit on fieldwork techniques, but mostly in in-house publications such as *Notes on Linguistics* or *Notes on Translation.* American dialectologists also wrote short pieces on fieldwork techniques in *American Speech*. These publications by SIL and in *American Speech* are listed in the references in Vaux et al. (2007:391–412).

own terms, rather than making descriptions conform to the mold of say, Latin or Sanskrit grammar, it remains true that the large majority of early writers were limited by known grammatical patterns.

Also, it is necessary to recognize that excellent fieldworkers might be more perceptive in their own fieldwork than when reading the result of the fieldwork of others. For example, if one is interested in Celtic, one will be impressed by the fieldwork of the Norwegian linguist Alf Sommerfelt (1892–1965), who wrote an indispensable study of a Breton dialect (Sommerfelt 1921, reedited 1978). However, this same Sommerfelt concluded, on the basis of misinterpretations of Strehlow's (see Section 3.6.2) writings on the Australian language Arrernte, that it was a 'primitive' language (Sommerfelt 1938), a conclusion rejected by modern Australianists (Moore 2008:287–288).

With the advent of formal fieldwork, first commissioned by museums and universities in the late nineteenth century, and then by granting agencies, one should hope that the average linguistic fieldworker will retain a healthy curiosity and perceptiveness, without having to be an exceptional individual and without giving up high standards in ethics as well as in linguistic adequacy.

References

Adelaar, Willem F. H. 1997. Las transiciones en la tradición gramatical hispanoamericana: historia de un modelo descriptivo. In *La descripción de las lenguas amerindias en la época colonial*, ed. by Klaus Zimmermann, 259–270. Frankfurt am Main: Vervuert; Madrid: Iberoamericana.

Adelaar, Willem F. H., with the collaboration of Pieter C. Muysken. 2004. *The Languages of the Andes*. (Cambridge Language Surveys.) Cambridge: Cambridge University Press.

von Adelung, Friedrich. 1976. *Catherinens der Grossen Verdienste um die Vergleichende Sprachenkunde*. Reprint of 1815 edition, ed. by H. Haarmann. Hamburg: Helmut Buske.

Armstrong, Robert G. 1964. *The Study of West African Languages.* Ibadan: Institute of African Studies.

Barbour, Stephen, and Patrick Stevenson. 1990. *Variation in German. A Critical Approach to German Sociolinguistics*. Cambridge: Cambridge University Press.

Barnum, Francis, S. J. 1893. *Life on the Alaska Mission*. Woodstock, MD: Woodstock College Press.

Bartlett, John R. 1851. [Vocabulary. Mimbreño language from Mancus Colorado, Chief of Coppermine Apaches.] Ms. BAE Nr. 121 and 1627, National Anthropological Archives, Smithsonian Institution, Washington, D.C.

Beach, D. M. 1938. *The Phonetics of the Hottentot Language*. Cambridge: W. Heffer & Sons.

Blake, Barry J. 2009. Book review of: Encountering Aboriginal Languages: Studies in the History of Australian Linguistics, ed. by William B. McGregor. *Australian Journal of Linguistics* 29(2):301–306.

Bloomfield, Leonard. 1962. *The Menomini Language*. New Haven, CT and London: Yale University Press.

Boas Yampolsky, Helene. 1958. Excerpts from the letter diary of Franz Boas on his first fieldtrip to the Northwest Coast. *International Journal of American linguistics* 24(4):312–320.

Bobrick, Benson. 1992. *East of the Sun. The Epic Conquest and Tragic History of Siberia*. New York: Poseidon Press.

Bouquiaux, Luc, and Jacqueline M. C. Thomas. 1976. *Enquête et description des langues à tradition orale. L'enquête de terrain et l'analyse grammaticale 1.; Approche linguistique 2; Approche thématique 3*. Paris: SELAF.

Bouquiaux, Luc, and Jacqueline M. C. Thomas, and James Robert. 1992. *Studying and Describing Unwritten Languages.* Dallas, TX: SIL International.

Bourke, John Gregory. 1980. *Vocabulary of the Apache or 'Inde language of Arizona & New Mexico, collected by John Gregory Bourke in the 1870s and 1880s*, ed. by Carole J. Condie. (Occasional Publications in Anthropology. Linguistics Series 7.) Greeley, CO: Museum of Anthropology, University of Northern Colorado.

Bridges, E. Lucas. 1951. *Uttermost Part of the Earth*. London: Hodder & Stoughton.

Bridges, Thomas. 1933. *Yamana-English. A Dictionary of the Speech of Tierra del Fuego*, ed. by Ferdinand Hestermann and Martin Gusinde. Mödling, Austria: Missionsdruckerei St. Gabriel. [Reprinted Buenos Aires: Zagier y Urruty, 1987]

Briggs, Lucy Therina. 1985. A Critical Survey of the Literature on the Aymara Language. In *South American Indian Languages. Retrospect and Prospect*, ed. by Harriet E. Manelis Klein and Louisa R. Stark, 546–594. Austin, TX: University of Texas Press.

Bright, William. 1967. Inventory of Descriptive Materials. In *Handbook of Middle American Indians*, Vol. 5, ed. by Norman A. McQuown, 9–62. Austin, TX: University of Texas Press.

Busbecq, Oghier Ghiselin de. 1968. *The Turkish letters of Ogier Ghiselin de Busbecq, Imperial Ambassador at Constantinople 1554–1562*. Trans. Edward Seymour Forster, 2nd ed. London: Oxford University Press.

Callaghan, Catherine A. 1991. Encounter with John P. Harrington. *Anthropological Linguistics* 33(4):350–356.

Calvo Pérez, Julio. 2004. Las Perífrasis Verbales en la Gramática Quechua de Diego González Holguín (1607). In *Missionary Linguistics/Lingüística Misionera: Selected Papers from the First International Conference on Missionary Linguistics,* Oslo, 13–16 March 2003, ed. by Otto Zwartjes and Even Hovdhaugen, 93–111. (Studies in the History of the Language Sciences 106.) Amsterdam and Philadelphia: John Benjamins.

Campbell, Lyle. 1997. *American Indian Languages. The Historical Linguistics of Native America*. New York and Oxford: Oxford University Press.

Canger, Una. 1990. Philology in America: Nahuatl: What Loanwords and the Early Descriptions of Nahuatl Show About Stress, Vowel Length, and Glottal Stop in Sixteenth Century Nahuatl and Spanish. In *Historical Linguistics and Philology*, ed. by Jacek Fisiak, 107–118. Berlin and New York: Mouton de Gruyter.

Canger, Una. 1994. Fieldwork and Field Methods. In *The Encyclopedia of Language and Linguistics*, Vol. 3, ed. by R. E. Asher, 1219–1220. Oxford: Pergamon Press.

Canger, Una. 1997. El Arte de Horacio Carochi. In *La descripción de las lenguas amerindias en la época colonial*, ed. by Klaus Zimmermann, 59–74. Frankfurt am Main: Vervuert; Madrid: Iberoamericana.

Capell, Arthur. 1945. *Methods and Materials for Recording Australian Languages.* Sydney, Australia: University of Sydney. Reprinted from Oceania 16(2):144–176 (1945).

Capell, Arthur. 1956. *A New Approach to Australian Linguistics*. Sydney: Oceania Linguistic Monographs.

Caraman, Philip. 1976. *The Lost Paradise. The Jesuit Republic in South America*. New York: The Seabury Press.

Carey, Hilary M. 2004. Lancelot Threlkeld and Missionary Linguistics in Australia to 1850. In *Missionary Linguistics/Lingüística Misionera: Selected Papers from the First International Conference on Missionary Linguistics,* Oslo, 13–16 March 2003, ed. by Otto Zwartjes and Even Hovdhaugen, 253–275. (Studies in the History of the Language Sciences 106.) Amsterdam and Philadelphia: John Benjamins.

Chambers, Jack K., and Peter Trudgill. 1980. *Dialectology*. (Cambridge Textbooks in Linguistics.) Cambridge: Cambridge University Press.

Chappell, Hilary. 2006. From Eurocentrism to Sinocentrism: The Case of Disposal Constructions in Sinitic Languages. In *Catching language: The Standing Challenge of Grammar Writing*, ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 441–485. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Childs, G. Tucker. 2007. The ethics of documentating dying languages: Lessons from the MDP and projections for the DKB. Gainesville, FL: 38th Annual Conference of African Linguistics. (ACAL), March 22–25.

Chomsky, Noam A. 1957. *Syntactic Structures*. The Hague: Mouton.

Christaller, Johann G. 1875. *A Grammar of the Asante and Fante Language, called Tshi.* Basel.

Collins, Beverley, and Inger M. Mees. 1999. *The Real Professor Higgins. The Life and Career of Daniel Jones*. Berlin and New York: Mouton de Gruyter.

Constantino, Ernesto Andres. 1959. *A Generative Grammar of a Dialect of Illocano* [sic]. Unpublished Ph.D. Dissertation, University of Indiana, Bloomington, IN.

Cooper, James Fenimore. 1826. *The Last of the Mohicans. A Narrative of 1757*. Philadelphia, PA: H. C. Carey & I. Lea.

Coseriu, Eugenio. 1967. Georg von der Gabelentz et la linguistique synchronique. *Word* 23(1–2–3):74–100.

Coupe, Alec R. 2007. *A Grammar of Mongsen Ao*. (Mouton Grammar Library 39.) Berlin and New York: Mouton de Gruyter.

Cowan, William, Michael K. Foster, and Konrad Koerner, eds. 1986. *New Perspectives in Language, Culture, and Personality: Proceedings of the Edward Sapir Centenary Conference*. (Ottawa, 1–3 October 1984). Amsterdam and Philadelphia: John Benjamins.

Crowley, Terry. 2007. *Field Linguistics. A Beginner's Guide*. Oxford: Oxford University Press.

Cyr, Danielle E. 1999. Past and New Directions for Fieldwork in Ethnolinguistics: The Case of Micmac (Northern Dialects). In *Theorizing the Americanist Tradition*, ed. by Lisa Philips Valentine and Regna Darnell, 268–281. Toronto: University of Toronto Press.

Daly, John P. 1973. *A Generative Syntax of Peñoles Mixtec*. Norman, OK: Summer Institute of Linguistics.

Darnell, Regna D. 1969. *The Development of American Anthropology, 1879 to 1920: From the Bureau of American Ethnology to Franz Boas*. Unpublished Ph.D. Dissertation in Anthropology, University of Pennsylvania, Philadelphia.

Darnell, Regna D. 1990. *Edward Sapir: Linguist, Anthropologist, Humanist*. Berkeley, CA: University of California Press.

Darnell, Regna D. 1999. Theorizing Americanist Anthropology: Continuities from the B.A.E. to the Boasians. In *Theorizing the Americanist Tradition*, ed. by Lisa Philips Valentine and Regna Darnell, 38–120. Toronto: University of Toronto Press.

Davis, Irvine. 1977. Native Americans as Linguists. *The Journal of the Linguistic Association of the Southwest* 2(2):47–53.

Dedenbach-Salazar Sáenz, Sabine. 2006. Quechua for Catherine the Great: José Joaquín Ávalos Chauca's Quechua Vocabulary. *International Journal of American Linguistics* 72(2):193–235.

Dixon, R. M. W. 1980. *The Languages of Australia*. (Cambridge Language Surveys.) Cambridge: Cambridge University Press.

Dixon, R. M. W. 1984. *Searching for Aboriginal Languages. Memoirs of a Field Worker*. Chicago, IL: University of Chicago Press.

Dixon, R. M. W. 2004. Ken Hale: An Appreciation. *Anthropological Linguistics* 46(3):342–345.

Dobrin, Lise M. 2009. SIL International and the disciplinary culture of linguistics. Introduction. *Language* 85(3):618–619.

Dobrin, Lise M., and Jeff Good. 2009. Practical language development: Whose mission? *Language* 85(3):618–629.

Doke, Clement M. 1926. *The Phonetics of the Zulu Language*. Johannesburg: University of Witwatersrand Press.

Doneux, Jean-Léonce. 2003. *Histoire de la linguistique africaine*. Aix-en-Provence: Publications de l'Université de Provence.

van Driem, George. 2001. *Languages of the Himalayas. An Ethnolinguistic Handbook of the Greater Himalayan Region*. (Handbook of Oriental Studies, Section 2, India, Volumes 10/1 and 10/2). Leiden: Brill.

Eliot, John. 1666. *The Indian Grammar Begun; or, an Essay to Bring the Indian language into Rules…* Cambridge, MA: Marmaduke Johnson. [Reprinted (1822) Massachusetts Historical Society Collections, 2nd series, Vol. 9, 274–312]

Emeneau, Murray B. 1940. Review of Lt. Col. D. L. R. Lorimer, the Burushaski language and E. O. Lorimer, Language hunting in the Karakoram. *Language* 16:351–357.

England, Nora C. 2007. The Influence of Mayan Speaking Linguists on the State of Mayan Linguistics. In *Endangered languages*, ed. by Peter K. Austin and Andrew Simpson, 93–111. Hamburg: Helmut Buske Verlag.

Epps, Patience, and Herb Ladley. 2009. Syntax, souls, or speakers? On SIL and community language development. *Language* 85(3):640–646.

Escalante, Fernando. 1990. Setting the record straight on Yaqui passives. *International Journal of American Linguistics* 56(2):289–292.

Evans, Nicholas. 2010. *Dying Words. Endangered Languages and What They Have to Tell Us*. Malden, MA: Wiley-Blackwell.

Evans-Pritchard, E. E. 1940. *The Nuer. A Description of the Modes of Livelihood and Political Institutions of a Nilotic people*. New York and Oxford: Oxford University Press.

Everett, Daniel L. 2004. Coherent Fieldwork. In *Linguistics Today – Facing a Greater Challenge*, ed. by Piet van Sterkenburg, 141–162. Amsterdam and Philadelphia: John Benjamins.

Fabian, Johannes. 1985. *Language on the Road. Notes on Swahili in Ttwo Nineteenth Century Travelogues*. (SUGIA-Beiheft 4). Hamburg: Helmut Buske.

Fabian, Johannes. 1986. *Language and Colonial Power. The Appropriation of Swahili in the Former Belgian Congo 1880–1938*. (African Studies Series 48.) Cambridge: Cambridge University Press.

Finck, Franz Nikolaus. 1899. *Die Araner Mundart. Ein Beitrag zur Erforschung des Westirischen*. Marburg: N. G.: Elwert'sche Verlagsbuch handling.

Firth, J. R. 1937. [1964.] *The Tongues of Men*. London: Oxford University Press.

Firth, J. R., et al. 1957. *Studies in Linguistic Analysis*. (Special Volume of the Philological Society.) Oxford: Basil Blackwell.

Foley, William A. 1986. *The Papuan languages of New Guinea*. (Cambridge Language Surveys.) Cambridge: Cambridge University Press.

Francis, W. N. 1983. *Dialectology. An Introduction*. London and New York: Longman.

Frantz, Donald. 1971. *Toward a generative grammar of Blackfoot*. Norman, OK: Summer Institute of Linguistics.

Fries Charles C. 1961. The Bloomfield 'School'. In *Trends in European and American Linguistics 1930–1960*, ed. by Christine Mohrmann, Alf Sommerfelt and Joshua Whatmough, 196–224. Utrecht & Antwerp: Spectrum

Gabelenz, Georg von der. 1891. *Die Sprachwissenschaft. Ihre Aufgaben, Methoden und bisherigen Ergebnisse*. Leipzig: Weigel.

García-Medall, Joaquín. 2004. Notas de lexicografía hispano-filipina: El Bocabulario de lengua bisaya, hiligueyna y haraya de la isla de Panay y Sugbu y para las demas islas, de fray Alonso de Méntrida, OSA, (ca. 1637). In *Missionary Linguistics/Lingüística Misionera: Selected Papers from the First International Conference on Missionary Linguistics,* Oslo, 13–16 March 2003, ed. by Otto Zwartjes and Even Hovdhaugen, 201–232. (Studies in the History of the Language Sciences 106.) Amsterdam and Philadelphia: John Benjamins.

Garvin, Paul L. 1964. *On Linguistic Method. Selected Papers.* The Hague: Mouton & Co.

George, Ken. 1993. Cornish. In *The Celtic languages*, ed. by Martin J. Ball and James Fife, 410–468. London and New York: Routledge.

Gilmour, Rachael. 2004. Colonization and Linguistic Representation. British Methodist Grammarian's Approaches to Xhosa (1834–1850). In *Missionary Linguistics/Lingüística Misionera: Selected Papers from the First International Conference on Missionary Linguistics,* Oslo, 13–16 March 2003, ed. by Otto Zwartjes and Even Hovdhaugen, 113–140. (Studies in the History of the Language Sciences 106.) Amsterdam and Philadelphia: John Benjamins.

Gilmour, Rachael. 2007. *Grammars of Colonialism. Representing Languages in Colonial South Africa*. New York: Palgrave MacMillan.

Gleason, Henry A. Jr. 1961. *An Introduction to Descriptive Linguistics*. New York: Holt, Rinehart & Winston.

Goddard, R. Ives. 1973. Publications on North American Indian Languages Issued by the Smithsonian Institution. *Themes in Linguistics: The 1970s*, ed. by Eric P. Hamp, 19–42. The Hague: Mouton.

Goddard, R. Ives. 1987. Leonard Bloomfield's descriptive and comparative studies of Algonquian. *Historiographia Linguistica* 14(1/2):179–217.

Goddard, R. Ives. 1996. The Description of the Native Languages of North America before Boas. *Handbook of North American Indians*, Vol. 17, ed. by Ives Goddard, 17–42. Washington D.C.: Smithsonian Institution.

Golla, Victor. 1991. John P. Harrington and his legacy. *Anthropological Linguistics* 33(4):337–349.

Grannier Rodriguez, Daniele Marcelle. 1997. La obra lingüística de Antonio Ruiz de Montoya. In *La descripción de las lenguas amerindias en la época colonial*, ed. by Klaus Zimmermann, 401–410. Frankfurt am Main: Vervuert; Madrid: Iberoamericana.

Gray, Edward G. 2000. Missionary Linguistics and the Description of "Exotic" Languages. In *History of the Language Sciences*, *Volume 18, Part 1. An International Handbook on the Evolution of the Study of Language from the Beginnings to the Present*, ed. by Sylvain Auroux, E. F. K. Koerner, Hans-Josef Niederehe, and Kees Versteegh, 929–937. Berlin: Walter de Gruyter.

Green, Jennifer, ed. 2001. 'Both Sides of the Bitumen': Ken Hale Remembering 1959. In *Forty Years on: Ken Hale and Australian Languages*, ed. by Jane Simpson, David Nash, Mary Laughren, Peter Austin and Barry Alpher, 29–43. Canberra: Pacific Linguistics

Gregersen, Edgar A. 1977. *Language in Africa. An Introductory Survey*. New York: Gordon & Breach.

Gregerson, Kenneth L. 1981. *A Study of Middle Vietnamese Phonology*. Dallas, TX: Summer Institute of Linguistics.

Grenoble, Lenore, and Lindsay J. Whaley. 2006. *Saving languages: An introduction to language revitalization*. Cambridge: Cambridge University Press.

Grierson, George A. 1903–1928. *Linguistic Survey of India*, Vols I-XI. Calcutta. [Reprint 1968, Delhi: Motilal Benarsidass.]

Gulya, Janos. 1974. Some Eighteenth Century Antecedents of Nineteenth Century Linguistics: The Discovery of Finno-Ugrian. In *Studies in the History of Linguistics*, ed. by Dell H. Hymes, 258–276. Bloomington, IN and London: Indiana University Press.

Haas, Mary R. 1953. Sapir and the training of anthropological linguists. *American Anthropologist* 55:447–450.

Haas, Mary R. 1968. Notes on a Chipewyan dialect. *International Journal of American Linguistics* 34(3):165–175.

Haas, Mary R. 1976. Boas, Sapir, and Bloomfield. In *American Indian Languages and American Linguistics*, ed. by Wallace L. Chafe, 59–69. Lisse: The Peter de Ridder Press.

Haeserijn, Esteban V. 1966. *Ensayo de la gramática del K'ekchi'*. Purulha, B.V.: Imprenta Suquinay.

Hair, P. E. H. 1967. *The Early Study of Nigerian Languages*. (West African Language Monographs 7.) Cambridge: Cambridge University Press.

Hale, Kenneth L. 1965. On the use of informants in fieldwork. *Canadian Journal of Linguistics* 10:108–119.

Hale, Kenneth L. 1972. Some Questions About Anthropological Linguistics: The Role of Native Knowledge. In *Reinventing Anthropology*, ed. by Dell Hymes, 382–397. New York: Pantheon Books.

Hale, Kenneth L. 1976. Theoretical Linguistics in Relation to American Indian Communities. In *American Indian Languages and American Linguistics*, ed. by Wallace L. Chafe, 35–50. Lisse: The Peter de Ridder Press.

Hale, Kenneth L. 1992a. On endangered languages and the safeguarding of diversity. *Language* 68(1):1–3.

Hale, Kenneth L. 1992b. Language endangerment and the human value of linguistic diversity. *Language* 68(1):35–42.

Hale, Sara Whitaker. 2001. Reminiscences of the Trip to Australia 1959–1961. In *Forty years on: Ken Hale and Australian Languages*, ed. by Jane Simpson, David Nash, Mary Laughren, Peter Austin and Barry Alpher, 19–28. Canberra: Pacific Linguistics

Hamp, Eric P. 1974. On Bonaparte and the Neogrammarians as Field workers. In *Studies in the History of Linguistics*, ed. by Dell H. Hymes, 390–433. Bloomington, IN and London: Indiana University Press.

Handman, Courtney. 2009. Language ideologies, endangered-language linguistics, and Christianization. *Language* 85(3):635–639.

Hanzeli, Victor Egon. 1969. *Missionary Linguistics in New France: A Study of Seventeenth- and Eighteenth-Century Descriptions of American Indian Languages*. The Hague: Mouton.

Harrington, John P. 1989. *Indian Tales from Picuris Pueblo*, ed. by Marta Weigle. Santa Fe, NM: Ancient City Press.

Hefley, James, and Marti Hefley. 1984. *Uncle Cam. The story of William Cameron Townsend, founder of the Wycliffe Bible Translators and the Summer Institute of Linguistics*. Huntington Beach, CA: Wycliffe Bible Translators.

Hercus, Luise. 2008. Listening to the Last Speakers. In *Encountering Aboriginal Languages: Studies in the History of Australian Linguistics*, ed. by William B. McGregor, 163–178. Canberra: Pacific Linguistics. Research School of Pacific and Asian Studies, The Australian National University.

Hill, Jane H. 1999. The Meaning of Writing and Text in a Changing Americanist Tradition. In *Theorizing the Americanist Tradition*, ed. by Lisa Philips Valentine and Regna Darnell, 180–194. Toronto: University of Toronto Press.

Hill, Jane H. 2006. Writing Culture in Grammar in the Americanist Tradition. In *Catching language: The Standing Challenge of Grammar Writing*, ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 609–628. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Hinton, Leanne. 1994. *Flutes of Fire. Essays on California Indian Languages*. Berkeley, CA: Heyday Books.

Hoijer, Harry. 1976. History of American Indian Linguistics. In *Native Languages of the Americas*, Vol. 1, ed. by Thomas A. Sebeok, 3–22. New York: Plenum Press.

Holm, John A. 1988. *Pidgins and Creoles. Volume I: Theory and Structure*. (Cambridge Language Surveys.) Cambridge: Cambridge University Press.

von Humboldt, Wilhelm. 1971. *Linguistic Variability and Intellectual Development*. Translated by George C. Buck and Frithjof A. Raven. Philadelphia: University of Pennsylvania Press.

Hymes, Dell H. 1963. Notes toward a history of linguistic anthropology. *Anthropological Linguistics* 5(1):59–103.

Hymes, Dell H. 1970. Linguistic Method in Ethnography: Its Development in the United States. In *Method and Theory in Linguistics*, ed. by Paul L. Garvin, 249–325. The Hague: Mouton & Co.

Hymes, Dell H. 1976. The Americanist Tradition. In *American Indian Languages and American Linguistics*, ed. by Wallace L. Chafe, 11–28. Lisse: The Peter de Ridder Press.

Hymes, Dell H. 2003. *Now I Know Only So Far. Essays in Ethnopoetics*. Lincoln and London: University of Nebraska Press.

Hymes, Dell H., and John Fought. 1981. *American Structuralism*. The Hague: Mouton.

Irvine, Judith T. 2008. Subjected words: African linguistics and the colonial encounter. *Language and Communication* 28:323–343.

Ivić, Milka. 1965. *Trends in Linguistics*. (Janua Linguarum, Series Minor No. XLII). The Hague: Mouton & Co.

Jakobson, Roman. 1958. Autobiographische Notizen von N. S. Trubetzkoy. In Nikolay S. Trubetzkoy, *Grundzüge der Phonologie*, 273–288. Göttingen: VandenHoek & Ruprecht.

Jäschke, H. A. 1881. *A Tibetan-English Dictionary. With Special Reference to the Prevailing Dialects. To Which is Added an English-Tibetan Vocabulary*. London and Henley: Routledge & Kegan Paul.

James, Carollyn. 1984. A field linguist who lived his life for his subjects. *Smithsonian* 15(1):153–174.

Jespersen, Otto. 1921. *Language. Its Nature, Development and Origin*. New York: W.W. Norton & Co.

Jetté, Jules, and Eliza Jones. 2000. *Koyukon Athabaskan Dictionary*. Fairbanks, Alaska: Alaska Native Language Center, University of Alaska.

Jones, Daniel, and Solomon T. Plaatje. 1916. *A Sechuana Reader in International Phonetic Orthography*. London: University of London Press.

Jungraithmayr, Hermann, and Wilhelm J. G. Möhlig, eds. 1983. *Lexicon der Afrikanistik*. Afrikanische Sprachen und Ihre Erforschung. Berlin: Dietrich Reimer.

Junquera Rubio, Carlos. 2005. *Fray José Pío Aza. Misionero, Geógrafo, Filólogo, Historiador, Etnólogo y Antropólogo en el Amazonas*. Pamplona, Spain: Ediciones Eunate.

Karttunen, Frances. 1991. A hundred years of Milpa Alta Nahuatl. *Estudios de cultura náhuatl* 21:271–308.

Karttunen, Frances. 2000. The linguistic career of Doña Luz Jiménez. *Estudios de cultura náhuatl* 30:267–274.

Kibrik, Aleksandr E. 2007. Collective Field Work. Advantages or Disadvantages? In *Perspectives on Grammar Writing*, ed. by Thomas E. Payne and David J. Weber, 25–44. (Benjamins Current Topics 11.) Amsterdam and Philadelphia: John Benjamins.

Klar, Kathryn A. 1991. "Precious beyond the power of money to buy": John P. Harrington's field-work with Rosario Cooper. *Anthropological Linguistics* 33(4):379–391.

Klein, Harriet E. Manelis, and Herbert S. Klein. 1978. The "Russian collection" of Amerindian languages in Spanish archives. *International Journal of American Linguistics* 44(2):137–144.

Kleinschmidt, Samuel. 1851. *Grammatik der grönländischen sprache, mit theilweisem einschluss des Labradordialects*. Berlin: G. Reimer.

Koerner, E. F. K. 1975. European Structuralism: Early Beginnings. In *Current Trends in Linguistics, Vol. 13. Historiography of Linguistics*, ed. by Thomas A. Sebeok, 717–827. The Hague: Mouton.

Koerner, E. F. K. 2004. Notes on Missionary Linguistics in North America. In *Missionary Linguistics/Lingüística Misionera: Selected Papers from the First International Conference on Missionary Linguistics,* Oslo, 13–16 March 2003, ed. by Otto Zwartjes and Even Hovdhaugen, 47–80. (Studies in the History of the Language Sciences 106.) Amsterdam and Philadelphia: John Benjamins.

Krauss, Michael E. 1973. *The Alaska Native Language Center Report, 1973*. Fairbanks, AK: University of Alaska.

Krauss, Michael E. 1992. The world's languages in crisis. *Language* 68(1):4–10.

Krauss, Michael E. 2006. A history of Eyak language documentation and study: Fredericae de Laguna in Memoriam. *Arctic Anthropology* 43(2):172–218.

Kroeber, Theodora. 1961. *Ishi in Two Worlds*. Berkeley, CA and Los Angeles, CA: University of California Press.

Kroeber, Theodora. 1964. *Ishi, Last of His Tribe*. Berkeley, CA: Parnassus Press.

Kummer, Werner. 1981. Malinche, Patron-Saint of Informants? In *A Festschrift for Native Speaker*, ed. by Florian Coulmas, 175–193. The Hague, Paris and New York: Mouton.

Ladefoged, Peter. 1992. Another view of endangered languages. *Language* 68(4):809–811.

Laird, Carobeth. 1975. *Encounter with an Angry God. Recollections of my Life with John Peabody Harrington*. Banning, CA: Malki Museum Press.

Landar, Herbert J. 1976a. Native North America. In *Native Languages of the Americas*, Vol. 1, ed. by Thomas A. Sebeok, 93–118. New York: Plenum Press.

Landar, Herbert J. 1976b. Historiography of Native Ibero-American Linguistics. In *Native Languages of the Americas*, Vol. 2, ed. by Thomas A. Sebeok, 185–206. New York: Plenum Press.

Lane, George S. 1945. Changes in emphasis in linguistics with particular reference to Paul and Bloomfield. *Studies in Philology* 42(3):465–483.

Laughlin, Robert M., and John B. Haviland. 1988. *The Great Tzotzil Dictionary of Santo Domingo Zinacantán, with Grammatical Analysis and Historical Commentary*. Three vols. (Smithsonian Contributions to Anthropology, No. 31.) Washington D.C.: Smithsonian Institution Press.

Laurent, Joseph. 1884. *New Familiar Abenakis and English Dialogues. The first ever published on the grammatical system*. Quebec: Leger Brousseau.

Lhuyd, Edward. 1707. *Archaeologia Brittannica*. [1969 reprint]. Menston, England: The Scolar Press Ltd.

Lindenfeld, Jacqueline. 1973. *Yaqui Syntax*. (University of California Publications in Linguistics 76.) Berkeley, CA and Los Angeles, CA: University of California Press.

Lorimer, David L. R. 1935a, 1935b, 1938. *The Burushaski Language*, Three vols. (Institutet for sammenlignende Kulturforskning, Serie B, 29.1, 2, 3.) Oslo: H, Aschehoug (W. Nygaard).

Lumwamu, François. 1973. *Essai de Morphosyntaxe Systématique des Parlers Kongo*. Paris: Klincksieck.

Machoni de Cerdeña, Antonio. 1877. *Arte y vocabulario de la lengua lule y tonocoté*. Buenos Aires: Pablo E. Coni. [orig. Madrid 1732]

Makkai, Adam. 1986. The Lexo-Centric Approach to Descriptive Linguistics. In *Language in Global Perspective: Papers in Honor of the 50th Anniversary of the Summer Institute of Linguistics 1935–1985*, ed. by Benjamin F. Elson, 47–61. Dallas, TX: Summer Institute of Linguistics.

Malinowski, Bronisław. 1989 [1967]. *A diary in the strict sense of the term*. Translated by Norbert Guterman. Stanford, CA: Stanford University Press.

Man, Edward H. 1923. *A Dictionary of the South Andaman (Âkà-Bêa) Language*. Bombay: British India Press.

Mandeville, Francois, and Ron Scollon. 2009. *This Is What They Say*. Vancouver and Toronto: Douglas & McIntyre, Seattle, WA: University of Washington Press.

Mannheim, Bruce. 1991. *The Language of the Inka Since the European Invasion*. Austin, TX: University of Texas Press.

Marsden, William. 1812. *A Grammar of the Malayan Language, with an Introduction and Praxis*. London: Cox & Baylis.

Martinet, André. 1945/1971. *La pronunciation du français contemporain, témoignages recueillis dans un camp d'officiers prisonniers*. [1st ed. 1945, 2nd ed. 1971] Paris and Geneva: Droz.

Maruyama, Toru. 2004. Linguistic Studies by Portuguese Jesuits in Sixteenth and Seventeenth Century Japan. In *Missionary Linguistics/Lingüística Misionera: Selected Papers from the First International Conference on Missionary Linguistics,* Oslo, 13–16 March 2003, ed. by Otto Zwartjes and Even Hovdhaugen, 140–177. (Studies in the History of the Language Sciences 106.) Amsterdam and Philadelphia: John Benjamins.

Masthay, Carl, ed. 1991. *Schmick's Mahican Dictionary*, with a Mahican historical phonology by David H. Pentland. (Memoirs of the American Philosophical Society, vol. 197.) Philadelphia: American Philosophical Society.

Mathiot, Madeleine. 1991. The reminiscences of Juan Dolores, an early O'odham linguist. *Anthropological Linguistics* 33(3):232–315.

Matthews, G. Hubert. 1965. *Hidatsa Syntax*. (Papers on Formal Linguistics 3.) The Hague: Mouton.

McGregor, William B., ed. 2008a. *Encountering Aboriginal Languages: Studies in the History of Australian Linguistics*. Canberra: Pacific Linguistics. Canberra: Research School of Pacific and Asian Studies, The Australian National University.

McGregor, William B. 2008b. Introduction. In *Encountering Aboriginal Languages: Studies in the History of Australian Linguistics*, ed. by William B. McGregor, 1–34. Canberra: Pacific Linguistics. Research School of Pacific and Asian Studies, The Australian National University.

McGregor, William B. 2008c. History of Fieldwork on Kimberley Languages. In *Encountering Aboriginal Languages: Studies in the History of Australian Linguistics*, ed. by William B. McGregor, 403–435. Canberra: Pacific Linguistics. Canberra: Research School of Pacific and Asian Studies, The Australian National University.

McGregor, William B., and Matti Miestamo. 2008. Nils M. Holmer's Research on Australian languages. In *Encountering Aboriginal Languages: Studies in the History of Australian Linguistics*, ed. by William B. McGregor, 219–250. Canberra: Pacific Linguistics. Research School of Pacific and Asian Studies, The Australian National University.

McQuown, Norman A. 1967. History of Studies in Middle American Linguistics. In *Handbook of Middle American Indians*, Vol. 5, ed. by Norman A. McQuown, 3–6. Austin, TX: University of Texas Press.

McQuown, Norman A. 1976. American Indian Linguistics in New Spain. In *American Indian Languages and American Linguistics*, ed. by Wallace L. Chafe, 105–127. Lisse: The Peter de Ridder Press.

Mead, Margaret. 1939. Native languages as fieldwork tools. *American Anthropologist* 41:181–205.

Mead, Margaret. 1972. *Blackberry Winter. My Earlier Years*. New York: William Morrow & Co.

Mead, Margaret. 1977. *Letters from the Field 1925–1975*. New York: Harper & Row.

Mead, Margaret, and Ruth L. Bunzel, eds. 1960. *The Golden Age of American Anthropology*. New York: George Braziller.

Medicine, Bea. 1999. Ella Cara Deloria: Early Lakota Ethnologist (Newly Discovered Novelist). In *Theorizing the Americanist Tradition*, ed. by Lisa Philips Valentine and Regna Darnell, 259–267. Toronto: University of Toronto Press

Miner, Kenneth L. 1974. John Eliot of Massachusetts and the beginnings of American linguistics. *Historiographia Linguistica* 1(2):169–183.

Mithun, Marianne. 1996. The Description of the Native Languages of North America: Boas and After. *Handbook of North American Indians*, Vol. 17, ed. by Ives Goddard, 43–63. Washington D.C.: Smithsonian Institution.

Moore, David Campbell. 2008. T. G. H. Strehlow and the linguistic Landscape of Australia 1930–1960. In *Encountering Aboriginal Languages: Studies in the History of Australian Linguistics*, ed. by William B. McGregor, 273–300. Canberra: Pacific Linguistics. Research School of Pacific and Asian Studies, The Australian National University.

Morice, Adrien Gabriel. 1904. Les langues dénées. *L'année Linguistique* 29:205–247. Paris: Librairie Klincksieck.

Mosel, Ulrike. 1987. *Inhalt und Aufbau deskriptiver Grammatiken (How to write a grammar)*. Institut für Sprachwissenschaft der Universität zu Köln, Arbeitspapier Nr. 4 (Neue Folge.) Cologne: Institut für Sprachwissenschaft, Universität Köln.

Mosel, Ulrike. 2006. Grammaticography: The Art and Craft of Writing Grammars. In *Catching language: The Standing Challenge of Grammar Writing*, ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 11–68. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Mühlhäusler, Peter. 1999. Exploring the missionary position. *Journal of Pidgin and Creole Languages* 14(2):339–346.

Mufwene, Salikoko S. 1998. The ecology of language: New imperatives in linguistics curricula. *Studies in the Linguistic Sciences* 28(2):135–145.

Múgica, Camilo. 1969. *Aprenda el guajiro. Gramática y vocabularios*. Riohacha, Guajira, Colombia: Vicario Apostólico de Riohacha.

Newman, Paul. 1998. We have seen the enemy and it is us: the endangered languages issue as a hopeless cause. *Studies in the Linguistic Sciences* 28(2):11–20.

Newmeyer, Frederick J. 1980. *Linguistic Theory in America. The First Quarter-Century of Transformational Generative Grammar*. New York: Academic.

Newmeyer, Frederick J. 1986. *The Politics of Linguistics*. Chicago, IL: University of Chicago Press.

O'Grady. Geoff. 2001. Tribute to Ken Hale: our 1960 Collaboration. In *Forty years on: Ken Hale and Australian Languages*, ed. by Jane Simpson, David Nash, Mary Laughren, Peter Austin and Barry Alpher, 227–230. Canberra: Pacific Linguistics.

Olesch, Reinhold. 1959. *Vocabularium Venedicum von Christian Hennig von Jessen*. Nachdruk besorgt von Reinhold Olesch. Cologne & Graz: Böhlau.

Olson, Kenneth S. 2009. SIL International: An emic view. *Language* 85(3):646–658.

Ostler, Nicholas. 2004. The Social Roots of Missionary Linguistics. In *Missionary Linguistics/ Lingüística Misionera: Selected Papers from the First International Conference on Missionary Linguistics,* Oslo, 13–16 March 2003, ed. by Otto Zwartjes and Even Hovdhaugen, 33–46. (Studies in the History of the Language Sciences 106.) Amsterdam and Philadelphia: John Benjamins.

Ostler, Nicholas. 2005. *Empires of the Word. A Language History of the World*. New York: HarperCollins.

Parks, Douglas R. 1999. George A. Dorsey, James R. Murie, and the textual Documentation of Skiri Pawnee. In *Theorizing the Americanist Tradition*, ed. by Lisa Philips Valentine and Regna Darnell, 227–244. Toronto: University of Toronto Press.

Pedersen, Holger. 1931/1962. *The Discovery of Language. Linguistic Science in the Nineteenth Century*. Bloomington, IN and London: Indiana University Press.

Pée, Willem. 1946. *Dialectatlas van West-Vlaanderen en Fransch-Vlaanderen*. (Reeks Nederlandsche Dialect-Atlassen 6.) Antwerp: De Sikkel.

Percival, W. Keith. 1974. Humboldt's Description of the Javanese verb. In *Studies in the History of Linguistics*, ed. by Dell H. Hymes, 380–389. Bloomington, IN and London: Indiana University Press.

Percival, W. Keith. 1981. *A Grammar of the Urbanised Toba-Batak of Medan*. (Pacific Linguistics Series B, No. 76.) Canberra: Department of Linguistics, Research School of Pacific Studies, The Australian National University.

Philips Valentine, Lisa, and Regna Darnell, eds. 1999. *Theorizing the Americanist Tradition*. Toronto: University of Toronto Press.

Pike, Kenneth L. 1947. *Phonemics. A Technique for Reducing Languages to Writing*. Ann Arbor, MI: University of Michigan Press.

Pike, Kenneth L. 1948. *Tone Languages. A Technique for Determining the Number and Type of Pitch Contrasts in a Language, With Studies in Tonemic Substitution and Fusion*. Ann Arbor, MI: University of Michigan Press.

Pike, Kenneth L. 1975. On Describing Languages. In *The Scope of American Linguistics. The First Golden Anniversary Symposium of the Linguistic Society of America*, ed. by Robert Austerlitz, 9–39. Lisse: The Peter de Ridder Press.

Poignant, Roslyn. 2004. *Professional Savages: Captive Lives and Western Spectacle*. New Haven, CT: Yale University Press.

Polański, Kazimierz, and James Allen Sehnert. 1967. *Polabian-English Dictionary*. The Hague: Mouton.

Porter, Joseph C. 1986. *Paper Medicine Man. John Gregory Bourke and his American West*. Norman and London: University of Oklahoma Press.

Postal, Paul M. 1979. *Some Syntactic Rules of Mohawk*. New York: Garland.

Ridruejo, Emilio. 2004. La primitiva tradición gramatical sobre el pampango. In *Missionary Linguistics/Lingüística Misionera: Selected Papers from the First International Conference on Missionary Linguistics,* Oslo, 13–16 March 2003, ed. by Otto Zwartjes and Even Hovdhaugen, 179–200. (Studies in the History of the Language Sciences 106.) Amsterdam and Philadelphia: John Benjamins.

Rigsby, Bruce. 1976. Discussion of the Americanist Tradition. In *American Indian Languages and American Linguistics*, ed. by Wallace L. Chafe, 29–31. Lisse: The Peter de Ridder Press.

Rodrigues, Aryon D. 1997. Descripción del tupinambá en el período colonial: el Arte de José de Anchieta. In *La descripción de las lenguas amerindias en la época colonial*, ed. by Klaus Zimmermann, 371–400. Frankfurt am Main: Vervuert; Madrid: Iberoamericana.

Rosing, Otto. 1951. Kleinschmidt Centennial II: Samuel Petrus Kleinschmidt. *International Journal of American Linguistics* 17:63–65.

Rousseau, André, ed. 1991. *Sur les traces de Busbecq et du Gotique*. (Collection UL3 Coordonnée par le Conseil Scientifique de l'Université Charles de Gaulle-Lille III, Travaux et Recherches. Lille: Presses Universitaires de Lille.

Rowe, John Howland. 1974. Sixteenth and Seventeenth Century Grammars. In *Studies in the History of Linguistics*, ed. by Dell H. Hymes, 361–379. Bloomington, IN and London: Indiana University Press.

Samarin, William J. 1967. *Field Linguistics. A Guide to Linguistic Field Work*. New York: Holt, Rinehart & Winston.

Sampson, Geoffrey. 1980. *Schools of Linguistics*. Stanford, CA: Stanford University Press.

Sapir, Edward, and Morris Swadesh. 1960. *Yana Dictionary*, ed. by Mary R. Haas. (University of California Publications in Linguistics 22.) Berkeley, CA and Los Angeles, CA: University of California Press.

Schleicher, August. 1871. *Laut- und Formenlehre der polabischen Sprache*. St. Petersburg: Commissionäre der Kaiserlichen Akademie der Wissenschaften, Eggers und Co., H. Schmitzdorff, J. Issakoff and A. Tcherkessoff.

Schuchardt, Hugo. 1979. *The ethnography of variation. Selected writings on pidgins and creoles*, edited and translated by Thomas L. Markey. Ann Arbor, MI: Karoma.

Schulze, Wolfgang. 2005. Grammars for East Caucasian. *Anthropological Linguistics* 47(3):321–352.

Seguin Anderson, Margaret, and Deanna Nyce. 1999. Nisga'a Studies and the Americanist tradition: Bringing First Nations Research and teaching into the Academy. In *Theorizing the Americanist Tradition*, ed. by Lisa Philips Valentine and Regna Darnell, 283–298. Toronto: University of Toronto Press.

Shaw, George Bernard. 1916/1941. *Pygmalion. A Romance in Five Acts*. London: Penguin Books.

Shuy, Roger W. 1990. A brief history of American sociolinguistics 1949–1989. *Historiographia Linguistica* 17(1/2): 183–209.

Singh, Udaya Narayana. 2006. The New linguistic Survey of India. Paper presented at South Asian Languages Analysis (SALA) Roundtable, Mysore, India Dec. 19–21, 2006.

Snyman, Jan W. 1970. *An Introduction to the !Xũ (!Kung) Language*. Cape Town: A. A. Balkema.

Sohn, Ho-min, and B. W. Bender. 1973. *A Ulithian Grammar.* (Pacific Linguistics C–27) Canberra: Research School of Pacific and Asian Studies, The Australian National University.

Sommerfelt, Alf. 1921. *Le Breton parlé à Saint-Pol-de Léon. Phonétique et morphologie*. Paris: Librairie Ancienne Honoré Champion, Édouard Champion.

Sommerfelt, Alf. 1938. *La langue et la société. Charactères sociaux d'une langue de type archaïque*. Oslo: Aschehoug.

Sommerfelt, Alf. 1978. *Le Breton parlé à Saint-Pol-de Léon. Phonétique et morphologie*. New edition by F. Falc'hun and Magne Oftedal. (Monographs in Celtic Studies from the University of Oslo 1.) Oslo-Bergen-Tromsø: Universitetsforlaget.

Stearns, MacDonald, Jr. 1978. *Crimean Gothic*. Analysis and Etymology of the Corpus. Saratoga, CA: Anma Libri.

Stocking, George W. Jr. 1992. *The Ethnographer's Magic and Other Essays in the History of Anthropology*. Madison, WI: The University of Wisconsin Press.

Stoll, David. 1982. *Fishers of Men or Founders of Empire? The Wycliffe Bible Translators in Latin America*. Cambridge, MA: Cultural Survival, Inc.

Sturtevant, William C. 2005. History of Research on the Native American Languages of the Southeast. In *Native Languages of the Southeastern United States*, ed. by Heather K. Hardy and Janine Scancarelli. Lincoln and London: University of Nebraska Press.

Sturtevant, William C., and Samuel Stanley. 1973. Linguistics in the Smithsonian. In *Themes in Linguistics: The 1970s*, ed. by Eric P. Hamp, 14–18. The Hague: Mouton.

Suárez, Jorge A. 1983. *The Mesoamerican Indian Languages*. (Cambridge Language Surveys.) Cambridge: Cambridge University Press.

Svelmoe, William L. 2009. 'We do not want to masquerade as linguists': A short history of SIL and the academy. *Language* 85(3):629–635.

Sweeney, Edwin R. 1998. *Mangas Coloradas. Chief of the Chiricahua Apaches*. Norman, OK: University of Oklahoma Press.

Sweet, Henry. 1877. *Handbook of Phonetics*. Oxford: Clarendon Press.

Tedlock, Dennis. 1983. *The Spoken Word and the Work of Interpretation*. Philadelphia: University of Pennsylvania Press.

Traill, Anthony. 1985. *Phonetic and Phonological Studies of !Xóõ Bushman*. Hamburg: Helmut Buske.

Trubetzkoy, Nikolay S. 1958. *Grundzüge der Phonologie*. Göttingen: VandenHoek & Ruprecht.

van der Tuuk, Hermanus N. 1864. *Tobasche Spraakkunst*. Vol 1. Klankstelsel. Amsterdam: Frederik Muller.

van der Tuuk, Hermanus N. 1867. *Tobasche Spraakkunst*. Vol. 2. De woorden als zinsdeelen. Amsterdam, The Netherlands: Depot van het Nederlandsche Bijbelgenootschap.

van der Tuuk, Hermanus N. 1971. *A grammar of Toba Batak*. Translated by Jeune Scott-Kemball, edited by A. Teeuw and R. Roolvink, with a foreword by A. Teeuw. (Koninklijk Instituut voor

Taal- Land- en Volkenkunde, Translation Series 13.) The Hague: Martinus Nijhoff. [Translation of van der Tuuk 1864–1867]

Vaux, Bert, Justin Cooper, and Emily Tucker. 2007. *Linguistic Field Methods*. Eugene, OR: Wipf & Stock.

Voegelin, Carl F. 1954a. A modern method for field work treatment of previously collected texts. *Journal of American Folklore* 67:15–20.

Voegelin, Carl F. 1954b. Multiple stage translation. *International Journal of American Linguistics* 20: 271–280.

Voegelin, Carl F. 1959a. Model-directed structuralization. *Anthropological Linguistics* 1(1):9–26.

Voegelin, Carl F. 1959b. The notion of arbitrariness in structural statement and restatement I: Eliciting. *International Journal of American Linguistics* 25(4):207–220.

Voegelin, Carl F. 1959c. Review of Eastern Ojibwa, by Leonard Bloomfield. *Language* 35(1): 109–125.

Voegelin, Carl F. 1960. Subsystem typology in linguistics. In *Man and Cultures: Selected papers*, ed. by Anthony F. C. Wallace, 202–206. (Fifth International Congress of Anthropological and Ethnological Sciences). Philadelphia: University of Pennsylvania Press.

Voegelin, Carl F., and Zellig S. Harris. 1945. Linguistics in ethnology. *Southwestern Journal of Anthropology* 1:455–465.

Voegelin, Carl F., and Zellig S. Harris. 1952. Training in anthropological linguistics. *American Anthropologist* 54:322–327.

Voegelin, Carl F., and Florence M. Robinett. 1954. Obtaining a linguistic sample. *International Journal of American Linguistics* 20:89–100.

Voegelin, Carl F., and Florence M. Voegelin. 1963. On the history of structuralizing in 20th century America. *Anthropological Linguistics* 5(1):12–37.

Wallis, Ethel E., and Mary A. Bennett. 1964. *Two Thousand Tongues to Go. The story of the Wycliffe Bible Translators*. New York: Harper & Row.

Wax, Rosalie H. 1971. *Doing Fieldwork. Warnings and Advice*. Chicago, IL: University of Chicago Press.

Westermann, Diedrich H., and Ida C. Ward. 1933. *Practical Phonetics for Students of African Languages*. London: Kegan Paul.

White, Leslie A. 1963. *The Ethnography and Ethnology of Franz Boas*. (Bulletin of the Texas Memorial Museum, Nr. 6.) Austin: Texas Memorial Museum.

Williams, Roger. 1643. *A Key into the Language of America: or, An help to the Language of the natives in that part of America, called New-England*. London: Gregory Dexter. [Republished, and edited with a critical Introduction, Notes, and Commentary by John J. Teunissen and Evelyn J. Hinz. Detroit, Wayne State University Press, 1973.]

Williamson, Kay. 1965. *A Grammar of the Kolokuma Dialect of Ijo*. Cambridge: Cambridge University Press.

Wmffre, Iwan. 1998. *Late Cornish*. (Languages of the World/Materials 135.) Munich and Newcastle: Lincom Europa.

Wonderly, William L., and Eugene A. Nida. 1963. Linguistics and Christian missions. *Anthropological Linguistics* 5(1):104–144.

Wrede, Ferdinand, et al. 1927–1956. *Deutscher Sprachatlas (DSA) auf Grund des Sprachatlas des deutschen Reichs*. Von Georg Wenker, begonnen v. Ferdinand Wrede, fortgesetzt v. Walther Mitzka u. Bernhard Martin. Marburg: Elwert.

Yengoyan, Aram A. 2001. On the Love of Languages. In *Forty years on: Ken Hale and Australian Languages*, ed. by Jane Simpson, David Nash, Mary Laughren, Peter Austin and Barry Alpher, 29–43. Canberra: Pacific Linguistics.

Zaefferer, Dietmar. 2006. Realizing Humboldt's dream: Cross-linguistic grammatography as database creation. In *Catching language: The Standing Challenge of Grammar Writing*, ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 113–135. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Zimmermann, Klaus, ed. 1997a. La *descripción de las lenguas amerindias en la época colonial*. Frankfurt am Main: Vervuert; Madrid: Iberoamericana.

Zimmermann, Klaus, 1997b. Introducción: apuntes para la historia de la linguistica de las lenguas amerindias. In *La descripción de las lenguas amerindias en la época colonial*, ed. by Klaus Zimmermann, 9–17. Frankfurt am Main: Vervuert; Madrid: Iberoamericana.

Zimmermann, Klaus. 2004. La construcción del objeto de la historiografía de la lingüística misionera. In *Missionary Linguistics/Lingüística Misionera: Selected Papers from the First International Conference on Missionary Linguistics,* Oslo, 13–16 March 2003, ed. by Otto Zwartjes and Even Hovdhaugen, 7–32. (Studies in the History of the Language Sciences 106.) Amsterdam and Philadelphia: John Benjamins.

Zwartjes, Otto, ed. 2000. *Las gramáticas misioneras de tradición hispánica (siglos XVI-XVII)*. Amsterdam and Atlanta: Rodopi.

Zwartjes, Otto, and Cristina Altman, eds. 2005. *Missionary Linguistics II/Lingüística Misionera II: Orthography and Phonology*. (Studies in the History of the Language Sciences 109.) Amsterdam and Philadelphia: John Benjamins.

Zwartjes, Otto, Ramón Arzápalo Marín, and Thomas C. Smith-Stark, eds. 2009. *Missionary Linguistics IV/Lingüística Misionera IV: Lexicography*. (Studies in the History of the Language Sciences 114.) Amsterdam and Philadelphia: John Benjamins.

Zwartjes, Otto, and Even Hovdhaugen, eds. 2004. *Missionary Linguistics /Lingüística Misionera: Selected Papers from the First International Conference on Missionary Linguistics,* Oslo, 13–16 March 2003. (Studies in the History of the Language Sciences 106.) Amsterdam and Philadelphia: John Benjamins.

Zwartjes, Otto, Gregory James, and Emilio Ridruejo, eds. 2007. *Missionary Linguistics III/ Lingüística Misionera III: Morphology and Syntax*. (Studies in the History of the Language Sciences 111.) Amsterdam and Philadelphia: John Benjamins.

Chapter 4
Choosing a Language

I know of one budding field linguist for example, who wanted to go to a tropical field location where there were no spiders! (Crowley 2007:61)

Even before fieldwork preparation (which will be dealt with in Chapter 5), it is necessary to choose a language to work on. However, the choice of language may well be out of the hands of the researcher. The language ultimately chosen for fieldwork may be suggested by an advisor or senior linguist (Section 4.1) or, due to special circumstances, a language community might request a researcher to conduct fieldwork on their language (Section 4.2). The situation where the researcher has the chance to choose a language by himself/herself, a case perhaps not as common as one might believe, will be treated in Section 4.3.

4.1 Another Outsider Chooses a Language for the Fieldworker

Many students do not have to select a language for fieldwork because an advisor or mentor can offer suggestions of possible languages to investigate. An engaging account of a professor choosing a language for his student is recounted in Dixon (1984:5–10). Krishnamurti (2007:56), in a short but very informative account of his impressive fieldwork career, explains how he was advised by his teacher T. Burrow to work on Konda, an undescribed Dravidian language of Central India. Oftentimes, the suggested language is of the same family or is closely related to a language the advisor is working on, and because of this the student can expect close guidance in data collection and analysis.[1]

From the perspective of the advisor, there may be interested students who are not the best choice to carry out fieldwork on a particular language. Some languages are so

[1]It is important to work on a language with input from linguists interested in and informed about the same language family. If a fieldworker wants to write a doctoral thesis based on fieldwork on a particular language, it is usually required that there be a professor or reader on the committee who specializes in that language family. In some institutions, the same requirement might hold for master's theses as well.

S.L. Chelliah and W.J. de Reuse, *Handbook of Descriptive Linguistic Fieldwork*,
DOI 10.1007/978-90-481-9026-3_4,

endangered that it might be better for an accomplished fieldworker to carry out the fieldwork, because accomplished fieldworkers can document a language faster and more efficiently, which is necessary with seriously endangered languages where time is of the essence. There are, of course, differing opinions on this point, but it is reasonable to suggest that the most urgent documentation efforts should not be put in the hands of M.A. or Ph.D. students, who will most likely be on their first fieldtrip, and who will typically have no experience in the field. No amount of preparation can stand in for experience. It is the responsibility of the professor to choose a language for a student in such a way that urgent documentation efforts will not be hampered or slowed down.

Some students – often, but not necessarily – those with a missionary background, have older relatives who worked on a language, and they are expected to continue work on this language. If the motivation is there, and the student has talent, this situation can work out. The student will have the appreciable advantage of introductions to the field and to native speakers.

4.2 The Language Community Chooses the Fieldworker

An optimal situation is one where the language community selects a fieldworker to work on their language. In these cases, the fieldworker cannot be accused of colonial or otherwise politically incorrect pursuits (Crowley 2007:80–81). In effect, the language community is in control of the fieldworker and of his/her fieldwork. Such situations usually concern language communities with a governing body that wants to document or preserve a language that is considered endangered. Well-known cases are that of Colette Grinevald's (formerly Craig's) involvement with the Rama language of Nicaragua (Craig 1992; Grinevald 2007), and that of David Wilkins' work with aboriginal languages of central Australia (Wilkins 1992).

In some instances, the aspiring fieldworker has relatives who speak the target language, or is married to someone who speaks the target language.[2] Fieldwork by family members is becoming more common as members of indigenous communities become interested in doing fieldwork on the heritage languages of their own communities. A good example of this is the case of Mayan linguists in Guatemala (England 2007), who rely as much on fieldwork as on introspection in their investigations.

4.3 The Fieldworker Chooses the Language

In this section, we do not imply, of course, that target language choice is something that should ever be a trivial or whimsical decision of one person, such as picking cherries from a tree, or chocolates out of a box.

[2] As we will see in Chapter 6, it is generally not a good idea to ask one's spouse to be one's fieldwork consultant; distant relatives are better. So if the only person the aspiring fieldworker wants to work with is his/her target-language-speaking spouse, s/he had better consider it carefully, and might try to find other speakers not related to him/her.

Ideally, a supervisor, or language community member will have a voice in a fieldworker's decision. However, in former times, when neither the scholarly world nor the community itself cared about language documentation, we are fortunate that someone, such as a "gentleman scholar", military, or colonial administrator (see Section 3.3), developed an interest in a particular language. There may still be cases where this type of "lone ranger" approach (Dwyer 2006:54) is the only one possible. It is not clear that this approach should automatically be considered a colonialist, anti-egalitarian, or unethical one. Consider the case of a student, perhaps someone with a missionary, Peace Corps or other developing country service background, who has gotten to know about a language isolate or even a language family for which there is no specializing professor to be found. S/he is, in effect, the fieldwork pioneer for that language or language family. There are still a few areas of the world where this could be the case: the Amazon, Papua New Guinea, and certain remote parts of the Pacific and Southeast Asia. True, it is increasingly rare to encounter such languages. Such a student is lucky, but his/her task will be harder: this enterprising student will have to work with a professor or advisor specializing in an unrelated language from the same geographical area.

Other students have taken a field methods course on a language no one in their educational institution is specializing in, but they liked the class so much that they want to do genuine fieldwork, generally on the same language or on a closely related language. If the language used in the fieldwork class is extremely well-studied, e.g. a language such as Farsi, or Turkish, or Swahili, it might not be a good idea to continue fieldwork on that language, unless the student knows of a particularly interesting topic regarding that language that needs to be elucidated through fieldwork. Then again, such a particular topic is often suggested by a professor. The original classroom consultant can help with community introductions (Crowley 2007:86–88) and thus this is a practical way to select a language to work on.

If the fieldworker finds himself/herself in a situation where s/he can choose the language, a variety of criteria (Section 4.3.1) and resources (Section 4.3.2) will influence language choice. Crowley (2007:57–61) is a helpful and anecdote-rich account of how to choose a language. A slightly different perspective is adopted in Bowern's guide to fieldwork (2008:126–127); she discusses choosing a field site, but not choosing a language.

4.3.1 Criteria Influencing Language Choice

The goals already described in Chapter 2 will, needless to say, influence language choice. These goals were:

- Primary goals: documentary and descriptive linguistics
- Secondary goals: religious instruction or the preservation of endangered languages
- Ancillary goals: non-comparative theoretical, comparative theoretical, dialectological or sociolinguistic, and the study of language, culture, and cognition

We will elaborate on selecting a language with the goal of description and language documentation using Dixon (2007) and Dwyer's (2006:51–52) useful advice on this topic. We use Dwyer's (2006:51) convenient criteria headings of Linguistic Diversity and/or Conservativeness (Section 4.3.1.1), Political Expediency (Section 4.3.1.2), Logistical Expediency (Section 4.3.1.3), and Interpersonal Expediency (Section 4.3.1.4), to which we will add a fifth one, Personal Expediency (Section 4.3.1.5).

4.3.1.1 Linguistic Diversity and/or Conservativeness

Because some languages are in graver danger of becoming extinct, it is preferable to choose an undescribed language or a language isolate, rather than a language belonging to a well-known family. And it is preferable to choose a language with previously unreported typological characteristics, or to choose a relatively conservative variety of the language.

Some languages are harder to learn (and therefore harder to describe) than others. Difficulty, however, is relative, depending partially on the fieldworker's language background. If the fieldworker is familiar with languages of family X, s/he will find it easier to describe a language of family X. If the fieldworker is familiar with an analytic language, he or s/he will find it easier to describe another analytic language. It is less clear to us, however, that experience with one sort of polysynthetic language, say an Eskimo one, is going to be much help with a polysynthetic language of another sort, say an Athabascan one, because there is quite a bit of typological variation within polysynthetic languages.[3]

The fieldworker can of course, choose a language because there is a particular aspect of its structure that interests him/her. This aspect can be phonetic, phonological, morphological, syntactic, semantic, pragmatic, or discourse-related. For example, my father's [de Reuse's] Dutch dialect is not particularly remarkable in terms of syntax and morphology, when compared to standard Dutch. However, it was worth studying phonetically and phonologically, because it exhibits several strange processes such as palatalization with unusual conditionings, coronal consonant harmony, vowel harmony, back unrounded vowels, and long and short syllabic nasals and laterals. None of these features are typical of Dutch or even of Germanic languages.

Dixon (2007) concludes by asking why it is that some linguists make a language seem dull and uninteresting, whereas other linguists can make a related language seem interesting and exciting. This tells one nothing about individual languages, but a lot about individual linguists. A brilliant fieldworker will always discover and bring out the interesting aspects of a language, no matter which language s/he chooses to work on, whereas a mediocre or uninspired linguist will make that same language look dull.

---

[3] See Section 11.3.2 for discussion of the term polysynthetic.

When the fieldworker needs to choose not only a particular language, but also a particular dialect or variety of a language to work on, s/he has to be aware of the political and language engineering consequences of that choice. The fieldworker might want to choose a dialect spoken, say, by 90% of the speech community, as opposed to one spoken by just 10% of the speech community, in the event that a choice is possible. Indeed, if the language is unwritten, (and assuming the community wants a writing system), it is almost certain that the particular dialect chosen will become the prestige dialect of the speech community. It serves the greatest number of that speech community if making the transition to the written dialect does not require that the majority learn a minority variety. Further discussion of the issue of language variation is in Section 5.3.

4.3.1.2 Political Expediency

Securing permissions is also a factor in selecting a language. Certain bureaucracies will make it so hard as to discourage fieldwork altogether. Obviously, there is no point in choosing a language for which the proper permissions are not going to be obtained. And obviously, permissions are going to be hard to obtain in politically unstable countries or in countries hostile to the fieldworker's home country.

But even among the indigenous communities of the U.S. and Canada, the degree of ease with which an American or a Canadian can obtain permissions varies enormously from reservation to reservation, or from reserve to reserve, and depends on the administration in power at a particular point in time. One cannot expect to get permissions more easily for an endangered language than for a non-endangered language. Some communities will reject requests for permission precisely because the proposal labels the language as "endangered", "moribund" or such terms, which they find offensive. More discussion regarding permissions is in Chapters 5 and 6.

Accessibility to areas varies constantly, and it is hard to generalize. For example, at present the field of Caucasian linguistics is severely restricted, and most fieldwork has to be done away from the Caucasus. Chukotka, in the Russian Far East, an area closed to outside research during the Cold War, was relatively open to outside researchers in the 1990s, but is now becoming more inaccessible. Similarly, there has been quite a bit of variation over the years as to which areas of China are open to outside linguistic research. Politically unstable areas of South Asia, such as Kashmir, Jharkand, the northeastern States of India, Nepal and Pakistan are also areas with difficult access. In Southeast Asia, Burma remains almost impossible for foreigners to enter for the purposes of up-country fieldwork.

4.3.1.3 Logistical Expediency

The remoteness of an area limits the ease with which the fieldworker can travel to or within the field site, or the number of villages that s/he can visit. It is impossible to estimate how much time it takes to travel to a relatively inaccessible area from looking at a map, and it is crucial to talk to someone who has actually traveled there.

The times of independently wealthy individuals who could travel anywhere with their own coolies or porters in Victorian comfort are over. Fieldworkers nowadays have to worry about the availability of funding for safe and efficient travel to a remote area. Availability of travel funding restricts language choices: funding may not be available for languages spoken in certain geographic areas or countries, in politically unstable or hostile countries; and funding may not be available for languages which are not perceived to be sufficiently endangered.

4.3.1.4 Interpersonal Expediency

It is always a good idea to consult with other fieldworkers and colleagues to make sure no one else has started doing the work the fieldworker wants to carry out. Generally, but not always, granting agencies will let one know if they cannot fund one's research because someone else is already doing it! But there are more pleasant ways of finding out that one's project should be modified.

We do not think that only one fieldworker should work on a particular language. The fieldworker does not own the field, after all. But it is professional courtesy to inquire if someone else is doing fieldwork on the language before starting, and if that is the case, then to negotiate with that other person, and to consider collaborating, dividing up the field of study (Crowley 2007:53). Bowern (2008:128) points out that once a language has been worked on, there is no "rule" that no other linguist can work on it.

However, duplication of effort should be avoided. It is often felt that two linguists in one field community are like two cooks in the kitchen, and two fieldworkers residing in the same location can lead to problematic and uncomfortable situations. The interaction described in Green (2001:33–34) of the young Ken Hale and the older Strehlow (see also Section 3.6.2), who were both fieldworkers interested in the Australian language Arrernte, is rather typical of the situations that can present themselves. It is preferable to study the same language in another location, with different native speakers, who most likely speak another dialect. The linguistics community, as well as the communities of speakers, usually benefit from diverse approaches.

4.3.1.5 Personal Expediency

Personal expediency concerns the restrictions related to the fieldworker him/herself, in the broadest sense: his/her religion, national or racial background, political views, gender, handicap, health, sexual orientation, his/her ability to deal with unusual climates, diets, long-distance and uncomfortable travel, or psychologically stressful situations, and his/her interests or tastes.

It is important to choose a language spoken in an area where one is not unbearably harassed or demeaned because of one's religion, national or racial background, political views, gender, handicap, or sexual orientation. As will be elaborated on in Chapter 5, one cannot expect people from other places to shares one's views or perspectives. As far as being harassed or demeaned is concerned, the fieldworker will need to develop a bit of a thick skin, but there are clear limits to this: no one should be expected to carry out fieldwork in a country plagued by epidemics, war, kidnappings, insurgencies, or terrorism. If the language is very interesting, the temptation is often there to take the risks. It is ultimately a personal decision of the fieldworker, a calculated risk (Bowern 2008:126), taken with the consideration of his/her family and academic advisors.

It is not necessary to be an Indiana Jones to carry out successful fieldwork but it is important to be able to rough it to some extent, as far as climate, diet, and transportation are concerned. We have known of vegetarians doing fieldwork with Eskimos or in China, so it can be done. One also needs to expect long distance travel or uncomfortable rides on small planes, all-terrain vehicles, snowmachines or skidoos, canoes, boats, jeeps, buses, trucks, or various animals.

Also, it is not necessary to have the steel nerves of a fighter pilot, but some psychological disorientation (known as culture shock) is to be expected. An aspiring fieldworker who is very sensitive to this might think twice about being a fieldworker in a culture very different from his/her own. The extent of culture shock varies, of course, according to the degree of difference from one's own culture. For an American or a European, culture shock will be extreme in certain areas of Asia, Africa, the Amazon, and the Pacific, but there will likely be much less culture shock when working on Native American reservations or with Australian aborigines.

Finally, fieldwork is also a matter of interests or personal taste: the aspiring fieldworker needs to have a feel or intuition that s/he is going to like the language, the culture, the people, and the way consultants carry out the work with him/her. Some fieldworkers have first been attracted by other cultural traits of the speakers of a language, such as local dress, music, or dance, before they get interested in the language.

Hyman (2001:29–32) thinks of fieldwork not so much as the state of being in an exotic location, but rather as "a state of mind". For him the qualities of a fieldworker are a love of discovery and a dedication to the whole language. But together with these qualities comes taste. As Hyman told us (p.c.) "I have never met a tone language I did not like."

4.3.2 Resources Influencing Language Choice

It is advisable to read about the language family or area for historical and typological background, and to read the existing anthropological literature as well. If possible, the aspiring fieldworker should read unpublished work on the language, just to make sure s/he will have something new or original to say. Background reading regarding philological study will be discussed in Section 5.2.

It is useful, while reading earlier sources to ask oneself the following questions: Are we talking about a distinct language, or is this just another name for a well-known language? Is the language a full-fledged language or is it a mutually intelligible variant of another language? Ethnologue (Lewis 2009) helps with answering some of these questions, but some of the perceptions about what is and what is not a language need to be taken with a grain of salt. See Hammarström's (2005) perceptive review of the previous edition of Ethnologue (Gordon 2005) on this issue.

In most cases, background reading will give the fieldworker a good idea of the general structure of the language, and the degree to which the language is endangered. Some excellent resources are listed below. They can be divided into Handbooks (Section 4.3.2.1), Language Survey volumes (Section 4.3.2.2), Bibliographies (Section 4.3.2.3), Atlases and Maps (Section 4.3.2.4), and Specialized Journals (Section 4.3.2.5).

4.3.2.1 Handbooks

Handbooks tend to have good introductions and language sketches or short grammars. Note that, while full of useful and reliable information, they tend to go quickly out of date regarding the most recent research. Handbook volumes of African languages are quite a bit older and therefore less useful in this regard.

Examples of handbooks are: Basset (1952): *Handbook of African Languages*, Vol. 1 on Berber; Bryan (1959): *Handbook of African Languages*, Vol 4 on Bantu; Derbyshire and Pullum (1986, 1990, 1991, 1998) *Handbook of Amazonian languages*, four volumes; Dixon and Blake (1979, 1981, 1983, 1991, 2000): *Handbook of Australian languages*, five volumes; Goddard (1996): *Handbook of North American Indians*, Volume 17: Languages; Tucker and Bryan (1956, 1966): *Handbook of African Languages*: Vol 3 on Northeast Africa, as well as Non-Bantu of South Africa; McQuown (1967): *Handbook of Middle American Languages*, Vol. 5, and Edmonson (1984): *Supplement to the Handbook of Middle American Languages*, Vol 2: Linguistics; and Westermann and Bryan (1952): *Handbook of African Languages*, Vol. 2 on West Africa. Van Driem (2001) is a rather personal but nonetheless informative, ethnolinguistic handbook of the languages of the Himalayas. It covers Tibeto-Burman, Indo-Aryan, and Burushaski, and adds lengthy digressions on Austronesian, Andamanese, Indo-European, Altaic, Yeniseian, and other Siberian.

4.3.2.2 Surveys

There exist two prestigious and reliable series of surveys. The *Cambridge Language Surveys* series tends to be organized in terms of geographical area, regardless of language family (except in the case of the Indian subcontinent), while the *Routledge Language Family Series* tends to be organized in terms of language families regardless of geography.

The *Cambridge Language Surveys* volumes which are likely to be of interest to linguistic fieldworkers cover the following languages:

- Amazonian indigenous languages: Dixon and Aikhenvald (1999)
- Andean indigenous languages, including languages of the surrounding areas not part of the Andes, thus all of South America except Amazonia, Venezuela, the Guyanas, eastern Paraguay, Uruguay and Brazil is covered: Adelaar and Muysken (2004)
- Australian indigenous languages: Dixon (1980, 2002)
- Chinese: Norman (1988)
- Dravidian: Krishnamurti (2003)
- Indo-Aryan: Masica (1991)
- Meso-American indigenous languages: Suárez (1983)
- North American indigenous languages: Mithun (1999)
- Papua New Guinean indigenous languages: Foley (1986)
- Pidgins and Creoles: Holm (1988, 1989)
- Languages of the (former) Soviet Union: Comrie (1981)

The *Routledge Language Family Series*[4] volumes which are likely to be of interest to linguistic fieldworkers cover the following language families:

- Austronesian: Himmelmann and Adelaar (2004)
- Bantu: Nurse and Philippson (2003)
- Dravidian: Steever (1998)
- Indo-Aryan: Cardona and Jain (2003)
- Iranian: Windfuhr (2009)
- Khoisan: Vossen (2009)
- Manchu-Tungusic: Vovin (2006)
- Mongolic: Janhunen (2003)
- Munda: Anderson (2008)
- Oceanic: Lynch et al. (2001)
- Semitic: Hetzron (1997)
- Sino-Tibetan: Thurgood and LaPolla (2003)
- Tai-Kadai: Diller et al. (2008)
- Turkic: Csató and Johanson (2006)
- Uralic: Abondolo (1998)

Other useful surveys, which are not part of any series, cover the following areas or languages:

- African languages: Heine et al. (1981); Heine and Nurse (2000); and Childs (2003)
- Australian indigenous languages: McConvell and Thieberger (2001) and Evans (2007)
- Caucasian: Klimov (1994)
- Indian languages: Abbi (2001:1–54)

[4] Over the years, this series has been called *Routledge Language Family Descriptions*, *Routledge Curzon Language Family Descriptions*, and *Curzon Language Family Descriptions*.

- Pidgins and Creoles: Kouwenberg and Singler (2008)
- Salish of North America: Czaykowska-Higgins and Kinkade (1998)
- South America indigenous languages: Klein and Stark (1985); Payne (1990); and Favre (1998)
- Uralic: Sinor (1988)
- Vanuatu languages: Lynch and Crowley (2001)

Miyaoka et al. (2007) is the most recent survey of the Pacific Rim, which includes, for purposes of this volume: pidgins and creoles of the Pacific, Mexico and Central America; the Pacific coast of South America; Tierra del Fuego; Australia; New Guinea; Malayo-Polynesian; China and Taiwan; Japanese and Ryukyuan; Nivkh; Ainu; Tungusic and Paleosiberian; Alaska; the North American Northwest coast; California; and the U.S. Southwest. While this volume is exemplary in providing detailed information on declining numbers of speakers, some of the information about work on these languages is somewhat out of date.

Introductory surveys which may be appropriate for beginning students who are not yet sure they want to be fieldworkers include Welmers (1973) and Gregersen (1977) for Africa, Goddard (2005) for the languages of East and Southeast Asia, Blake (1981) for Australia, and Lynch (1998) for Australia and the Pacific.

4.3.2.3 Bibliographies

References to bibliographies can be found in the handbooks and surveys above, and if the fieldworker can sift through potentially spurious or inaccurate references, there are many bibliographies on the Internet. As for handbooks and surveys, it is important to keep in mind the compilation date of the bibliography. As model printed bibliographies, we recommend Carrington (1996) for New Guinea, and Carrington and Triffit (1999) for Australia. Bibliographies are most useful, of course, if they are annotated.

4.3.2.4 Atlases and Maps

Of course the handbooks and surveys mentioned will also contain maps, but sometimes the fieldworker might consult larger atlases of languages just to see exactly where s/he would like to go, and what languages are spoken in the surrounding areas.

Ethnologue (Lewis 2009) and Asher and Moseley (2007) are the most comprehensive works, but the reliability of the map information is variable.

Moseley (2010) is an atlas for endangered languages; it is always instructive to compare this atlas with an earlier edition such as Wurm (2001), to remind us to which extent the concepts of “language” and “endangered” are politically charged.

Wurm et al. (1996) is an extremely detailed atlas covering intercultural communication in the Pacific, Asia, and the Americas.

For some countries or areas, there exist detailed language atlases, such as Heine and Möhlig (1980) for Kenya, and Taber (1996) for the Maluku area of Indonesia.

If the fieldworker wants to see where certain typological features occur with the greatest concentration, s/he can refer to the unique atlas of typological features by Haspelmath et al. (2005).

4.3.2.5 Specialized Journals

There are of course many journals that have areal concentrations, and these should be read for the latest research in a particular area or family. The titles of these journals can easily be found in the handbooks and surveys above.

References

Abbi, Anvita. 2001. *A Manual of Linguistic Field Work and Indian Language Structures.* (Lincom Handbooks in Linguistics 17.) Munich: Lincom Europa.

Abondolo, Daniel, ed. 1998. *The Uralic Languages*. (Routledge Language Family Descriptions.) London and New York: Routledge.

Adelaar, Willem F. H., with the collaboration of Pieter C. Muysken. 2004. *The Languages of the Andes*. (Cambridge Language Surveys.) Cambridge: Cambridge University Press.

Anderson, Gregory D. S., ed. 2008. *The Munda Languages*. (Routledge Language Family Series.) London and New York: Routledge.

Asher, R. E., and Christopher Moseley, eds. 2007. *Atlas of the World's Languages*. 2nd edition. New York: Routledge.

Basset, André. 1952. *La langue berbère*. (Handbook of African Languages Part 1.) London, New York, and Toronto: Oxford University Press.

Blake, Barry J. 1981. *Australian Aboriginal Languages*. London, Sydney, and Melbourne: Angus & Robertson.

Bowern, Claire. 2008. *Linguistic Fieldwork. A Practical Guide*. New York: Palgrave MacMillan.

Bryan, Margaret A. 1959. *The Bantu Languages of Africa*. (Handbook of African Languages Part 4.) London, New York, and Toronto: Oxford University Press.

Cardona, George, and Dhanesh Jain, eds. 2003. *The Indo-Aryan Languages*. (Routledge Language Family Series.) London and New York: Routledge.

Carrington, Lois. 1996. *A Linguistic Bibliography of the New Guinea Area*. (Pacific Linguistics D-90.) Canberra: Research School of Pacific and Asian Studies, The Australian National University.

Carrington, Lois, and Geraldine Triffit. 1999. *OZBIB: A Linguistic Bibliography of Aboriginal Australia and the Torres Strait Islands*. (Pacific Linguistics D-92.) Canberra: Research School of Pacific and Asian Studies, The Australian National University.

Childs, G. Tucker. 2003. *An Introduction to African Languages*. John Benjamins: Amsterdam and Philadelphia.

Comrie, Bernard. 1981. *The Languages of the Soviet Union*. (Cambridge Language Surveys.) Cambridge: Cambridge University Press.

Craig, Colette G. 1992. A constitutional response to language endangerment: The case of Nicaragua. *Language* 68(1):17–24.

Crowley, Terry. 2007. *Field Linguistics. A Beginner's Guide*. Oxford: Oxford University Press.

Csató, Éva Á., and Lars Johanson, eds. 2006. *The Turkic Languages*. (Routledge Language Family Series.) London and New York: Routledge.

Czaykowska-Higgins, Ewa, and M. Dale Kinkade. 1998. Salish Languages and Linguistics. In *Salish Languages and Linguistics. Theoretical and Descriptive Perspectives*, ed. by Ewa Czaykowska-Higgins and M. Dale Kinkade, 1–68. Berlin and New York: Mouton de Gruyter.

Derbyshire, Desmond C., and Geoffrey K. Pullum, eds. 1986, 1990, 1991, 1998. *Handbook of Amazonian Languages* (4 volumes). Berlin and New York: Mouton de Gruyter.

Diller, Anthony, Jerry Edmondson, and Yongxian Luo, eds. 2008. *The Tai-Kadai Languages*. (Routledge Language Family Series.) London and New York: Routledge.

Dixon, R. M. W. 1980. *The Languages of Australia*. (Cambridge Language Surveys.) Cambridge: Cambridge University Press.

Dixon, R. M. W. 1984. *Searching for Aboriginal Languages. Memoirs of a Field Worker*. Chicago, IL: University of Chicago Press.

Dixon, R. M. W. 2002. *Australian Languages. Their nature and Development*. (Cambridge Language Surveys.) Cambridge: Cambridge University Press.

Dixon, R. M. W. 2007. Field Linguistics: a minor manual. *Sprachtypologie und Universalienforschung* (Focus on Linguistic Fieldwork, ed. by Alexandra Y. Aikhenvald) 60(1):12–31.

Dixon, R. M. W., and Alexandra Y. Aikhenvald, eds. 1999. *The Amazonian Languages*. (Cambridge Language Surveys.) Cambridge: Cambridge University Press.

Dixon, R. M. W., and Barry J. Blake, eds. 1979, 1981, 1983, 1991, 2000. *Handbook of Australian Languages*. (5 volumes.) Canberra: Australian National University Press (vols. 1,2,3); Oxford: Oxford University Press (vols. 4, 5).

van Driem, George. 2001. *Languages of the Himalayas. An Ethnolinguistic Handbook of the Greater Himalayan Region*. (Handbook of Oriental Studies, Section 2, India, Volumes 10/1 and 10/2). Leiden: Brill.

Dwyer, Arienne. 2006. Ethics and practicalities of cooperative fieldwork and analysis. In *Essentials of Language Documentation*, ed. by Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel, 31–66. (Trends in Linguistics. Studies and Monographs 178.) Berlin and New York: Mouton de Gruyter. Berlin: Mouton de Gruyter.

Edmonson, Munro S., ed. 1984. Linguistics. *Supplement to the Handbook of Middle American Indians*, Vol 2. Austin, TX: University of Texas Press.

England, Nora C. 2007. The influence of Mayan speaking linguists on the state of Mayan linguistics. In *Endangered Languages*, ed. by Peter K. Austin and Andrew Simpson, 93–111. Hamburg: Helmut Buske Verlag.

Evans, Nicholas. 2007. Warramurrugunji undone: Australian languages in the 51st Millenium. In *Endangered Languages*, ed. by Peter K. Austin and Andrew Simpson, 19–44. Hamburg: Helmut Buske Verlag.

Favre, Alain. 1998. *Manual de las lenguas indígenas sudamericanas*. (2 volumes.) Munich and Newcastle: LINCOM Europa.

Foley, William A. 1986. *The Papuan Languages of New Guinea*. (Cambridge Language Surveys.) Cambridge: Cambridge University Press.

Goddard, Cliff. 2005. *The Languages of East and Southeast Asia. An Introduction.* Oxford: Oxford University Press.

Goddard, R. Ives, ed. 1996. *Handbook of North American Indians*, Vol. 17: Languages. Washington D.C.: Smithsonian Institution.

Gordon, Raymond, ed. 2005. *Ethnologue: Languages of the World*, 15th edition. Dallas, TX: SIL International.

Green, Jennifer, ed. 2001. 'Both sides of the bitumen': Ken Hale remembering 1959. In *Forty years on: Ken Hale and Australian Languages*, ed. by Jane Simpson, David Nash, Mary Laughren, Peter Austin and Barry Alpher, 29–43. Canberra: Pacific Linguistics.

Gregersen, Edgar A. 1977. *Language in Africa. An Introductory Survey*. New York: Gordon & Breach.

Grinevald, Colette. 2007. Linguistic Fieldwork among Speakers of Endangered Languages. In *The Vanishing Languages of the Pacific Rim*, ed. by Osahito Miyaoka, Osamu Sakiyama, and Michael Krauss, 35–76. Oxford: Oxford University Press.

Hammarström, Harald 2005. Review of: Ethnologue: Languages of the World, 15th Edition. *The Linguist List* 16.2637, Sept. 12, 2005. Online: http://www.ling.ed.ac.uk/linguist/issues/16/16–2637.html

Haspelmath, Martin, Matthew S. Dryer, David Gil, and Bernard Comrie, eds. 2005. *The World Atlas of Language Structures*. Oxford and New York: Oxford University Press.

Heine, Bernd, and Wilhelm W. J. Möhlig, eds. 1980. *Language and Dialect Atlas of Kenya*. Berlin: Reimer.

Heine, Bernd, and Derek Nurse. 2000. *African Languages*. An Introduction. Cambridge: Cambridge University Press.

Heine, Bernd, Thilo C. Schadeberg, and Ekkehard Wolff, eds. 1981. *Die Sprachen Afrikas*. Hamburg: Helmut Buske.

Hetzron, Robert. 1997. *The Semitic Languages*. (Routledge Language Family Descriptions.) London and New York: Routledge.

Himmelmann, Nikolaus P., and Sander Adelaar, eds. 2004. *The Austronesian Languages of Asia and Madagascar*. (Routledge Language Family Descriptions.) London and New York: Routledge.

Holm, John A. 1988. *Pidgins and Creoles. Volume I: Theory and Structure*. (Cambridge Language Surveys.) Cambridge: Cambridge University Press.

Holm, John A. 1989. *Pidgins and Creoles. Volume II: Reference Survey*. (Cambridge Language Surveys.) Cambridge: Cambridge University Press.

Hyman, Larry M. 2001. Fieldwork as a state of mind. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 15–33. Cambridge: Cambridge University Press.

Janhunen, Juha, ed. 2003. *The Mongolic Languages*. (Routledge Language Family Series.) London and New York: Routledge.

Klein, Harriet E. Manelis, and Louisa R. Stark, eds. 1985. *South American Indian Languages. Retrospect and Prospect*. Austin, TX: University of Texas Press.

Klimov, Georgiy A. 1994. *Einführung in die kaukasische Sprachwissenschaft*. Translated from the Russian by Jost Gippert. Hamburg: Helmut Buske.

Kouwenberg, Silvia, and John Victor Singler. 2008. *The Handbook of Pidgin and Creole Studies*. Malden, MA: Wiley-Blackwell.

Krishnamurti, Bhadriraju. 2003. *The Dravidian Languages*. (Cambridge Language Surveys.) Cambridge: Cambridge University Press.

Krishnamurti, Bhadriraju. 2007. Fieldwork on Konda, a Dravidian language. *Sprachtypologie und Universalienforschung* (Focus on Linguistic Fieldwork, ed. by Alexandra Aikhenvald) 60(1):56–66.

Lewis, M. Paul, ed. 2009. *Ethnologue: Languages of the World*, 16th edition. Dallas, TX: SIL International. Online version: http://www.ethnologue.com.

Lynch, John. 1998. *Pacific Languages: An Introduction*. Honolulu: University of Hawai'i Press.

Lynch, John, and Terry Crowley. 2001. *Languages of Vanuatu: A New Survey and Bibliography*. Canberra: Pacific Linguistics.

Lynch, John, Malcolm Ross, and Terry Crowley, eds. 2001. *The Oceanic Languages*. (Routledge Language Family Series.) London and New York: Routledge.

Masica, Colin. 1991. *The Indo-Aryan Languages*. (Cambridge Language Surveys.) Cambridge: Cambridge University Press.

McConvell, Patrick, and Nicholas Thieberger. 2001. *State of Indigenous languages in Australia – 2001*. Australia State of the Environment Second Technical Paper Series (Natural and Cultural Heritage). Canberra: Department of the Environment and Heritage.

McQuown, Norman A., ed. 1967. *Linguistics. Handbook of Middle American Indians*, Vol. 5. Austin, TX: University of Texas Press.

Mithun, Marianne. 1999. *The Languages of Native North America*. (Cambridge Language Surveys.) Cambridge: Cambridge University Press.

Miyaoka, Osahito, Osamu Sakiyama, and Michael Krauss, eds. 2007. *The Vanishing Languages of the Pacific Rim*. Oxford: Oxford University Press.

Moseley, Christopher. 2010. *Atlas of the World's Languages in Danger*. 3rd edition. Paris: UNESCO.

Norman, Jerry. 1988. *Chinese*. (Cambridge Language Surveys.) Cambridge: Cambridge University Press.

Nurse, Derek, and Gérard Philippson, eds. 2003. *The Bantu Languages*. (Routledge Language Family Series.) London and New York: Routledge.

Payne, Doris L., ed. 1990. *Amazonian Linguistics. Studies in Lowland South American Languages*. Austin, TX: University of Texas Press.

Sinor, Denis, ed. 1988. *The Uralic Languages: Description, History, and Foreign Influences*. Leiden and New York: Brill.

Steever, Sanford B., ed. 1998. *The Dravidian Languages*. (Routledge Language Family Descriptions.) London and New York: Routledge.

Suárez, Jorge A. 1983. *The Mesoamerican Indian Languages*. (Cambridge Language Surveys.) Cambridge: Cambridge University Press.

Taber, Mark H., ed. 1996. *Atlas bahasa tanah Maluku (Maluku Language Atlas).* Ambon: Pusat Pengkajian dan Pengembangan Maluku, Universitas Pattimura, and Summer Institute of Linguistics.

Thurgood, Graham, and Randy J. LaPolla, eds. 2003. *The Sino-Tibetan Languages*. (Routledge Language Family Descriptions.) London and New York: Routledge.

Tucker, A. N., and Margaret A. Bryan. 1956. *The Non-Bantu Languages of North-Eastern Africa*. (Handbook of African Languages Part 3.) London, New York, and Toronto: Oxford University Press.

Tucker, A. N., and Margaret A. Bryan. 1966. *Linguistic Analyses: The Non-Bantu Languages of North-Eastern Africa*. (Handbook of African Languages Part 5.) London, New York, and Toronto: Oxford University Press.

Vossen, Rainer, ed. 2009. *The Khoesan Languages*. (Routledge Language Family Series.) London and New York: Routledge.

Vovin, Alexander, ed. 2006. *The Manchu-Tungusic languages*. (Curzon Language Family Series.) Richmond: Curzon Press.

Welmers, William E. 1973. *African Language Structures*. Berkeley, CA and Los Angeles, CA: University of California Press.

Westermann, Diedrich H., and Margaret A. Bryan. 1952. *Languages of West Africa*. (Handbook of African Languages Part 2.) London, New York, and Toronto: Oxford University Press.

Wilkins, David. 1992. Linguistic research under Aboriginal control: a personal account of fieldwork in Australia. *Australian Journal of Linguistics* 12(1):171–200.

Windfuhr, Gernot, ed. 2009. *The Iranian Languages*. (Routledge Language Family Series.) London and New York: Routledge.

Wurm, Stephen A. 2001. *Atlas of the World's Languages in Danger of Disappearing*. 2nd edition. Paris: UNESCO.

Wurm, Stephen A., Peter Mühlhäusler, and Darrell T. Tryon, eds. 1996. *Atlas of Languages of Intercultural Communication in the Pacific, Asia, and the Americas*. (Trends in Linguistics Documentation 13.) Berlin and New York: Mouton de Gruyter.

Chapter 5
Field Preparation: Philological, Practical, and Psychological

5.1 Introduction

Having decided to undertake descriptive linguistic fieldwork, the researcher must tackle the complex preparations necessary for the trip. First, in addition to general typological study (see Chapter 11), a fieldworker must read materials specific to the language and culture being investigated. Second, a fieldworker must know what to expect from the field situation and must consider how his or her individual personality will respond to and engage with the pressures of the field. Third, a fieldworker must deal with practical arrangements before leaving for the field, including seeking funding, making contacts with a community of speakers, and purchasing and learning to use the right equipment. Finally, the fieldworker must obtain appropriate permissions from local authorities, be they from the fieldworker's home institution, funding agencies, and/or central, local, or tribal governments at the field site.

5.2 Philological Preparation

It is advisable that a fieldworker prepare for linguistic fieldwork by extracting relevant information from everything that has already been written about the target language. This kind of study, called philology, has been defined by Goddard (1976:72) as "part of the discipline of linguistics that is concerned with getting from texts and other recorded attestations of languages systematic information that is not directly conveyed by such records as they stand." Goddard distinguishes a branch of philology that is concerned with gathering information on cultures, from a branch of philology concerned with gathering information about the languages themselves. As descriptive fieldworkers, we are primarily concerned with the second branch. But even the second definition implies a sort of linguistic study – i.e. the study of documents or records – that has nothing to do with descriptive fieldwork. In the next section we will show how philology actually is relevant to fieldwork.

S.L. Chelliah and W.J. de Reuse, *Handbook of Descriptive Linguistic Fieldwork*,
DOI 10.1007/978-90-481-9026-3_5,

5.2.1 Why Philological Preparation Matters

Why is philology relevant to fieldwork? The connection is as follows. A substantial portion of records of lesser-known languages, or generally unwritten languages, or endangered languages, consists in word lists, field notes, unpublished grammars, dictionaries, text collections, and religious materials written by a variety of people. Many of these people were amateurs, gentleman scholars, or missionaries, as we discussed in Sections 3.2 and 3.3.[1] Fewer and fewer languages of the world have absolutely nothing written about them, so part of fieldwork preparation is to read and use the old records, if they exist.

There also exist archival sound recordings of languages, of course, and these should be studied philologically as well (Bowman 1959; Levitt 1989; Golla 1995:145–146, Tuttle 2003). However, sound recordings are more recent and far less numerous than documents on paper. Therefore we will concentrate on philological uses of written documents in the rest of this section.

Some fieldworkers will balk at having to read the older material. The stuff can be hard to find, hard to read, in funny spellings, written in strange languages, poorly archived, and may be of overall poor quality (Goddard 1976:74). The fieldworker might be tempted to think that whatever is in these old documents can be elicited and analyzed faster, better, and more accurately by him-/herself. Even so, there are three reasons why philological preparation is necessary.

First, the fieldworker can learn from the previous recorder's spelling, archiving, and analytical mistakes, and can learn to distinguish an insightful grammar or text collection from a mediocre one (Bowern 2008:190–192). (It is sobering to think that one's own field notes might be as incomprehensible to researchers 200 years from now as the messy, perversely misanalysed data in documents we are faced with from 200 years ago. This is, of course, one of the reasons why archiving one's field notes is important. Modern archiving techniques will be discussed in Chapter 7.)

Second, even if the documents are of very poor quality, it is the responsibility of the fieldworker to see if there is anything in them that might be of interest, such as old words or forms which s/he might not be able to elicit in the modern language. This is especially true when documenting or describing a severely endangered language. It is necessary to add philological information to one's own documentation to make it as comprehensive as possible. One superb example of fieldwork complemented by extensive philological work is the Nahuatl dictionary by Karttunen (1983), which is based on sixteenth, seventeenth, and eighteenth century dictionaries and documents, as well as on dictionaries based on twentieth century fieldwork. Another is the account of Eyak language documentation, by Krauss (2006), which is a discussion of everything ever written in or about the Eyak language of southern Alaska, culminating in Krauss's own fieldwork.

[1] There is some overlap between the discussion in this section and the historical presentation in Chapter 3. For example, the first book-length study of missionary linguistics, focusing on seventeenth and eighteenth century New France (Hanzeli 1969), is both a history of these documents and of their writers, as well as a philological study.

Third, and closely related to the second point, the fieldworker needs to study the old documents to see what is missing, so that s/he can re-elicit and/or build on the old documents. The restoration and interpretation of texts is a particularly useful area in which fieldworkers collaborate with philologists (Goddard 1976:85).

5.2.2 *Advice on Philological Work*

The foundational essays on Native American language philology (Goddard 1976; Haas 1975) were written by fieldworkers. Thieberger's practical manual (1995), intended to teach to the non-linguist the basic methods of philology of Australian aboriginal languages with a view to revitalization, was also written by a fieldworker.[2] Besides these works, there is not much guidance for the philology of documents in little-known or endangered languages. The craft of philology is best learned by looking at good examples of different kinds of philology. For more philosophical approaches to philology, and the relationship between fieldwork, introspection, and philology, we recommend Anttila (1979) and Ehlich (1981).

In the rest of this section, we will give some philological advice to the fieldworker, with selected references to good philological practice.

5.2.2.1 The Basic Rules

The most basic rule of philology is this: the philologist should always distinguish clearly and unambiguously between the original document itself and his/her own elicited additions, corrections, and editions to the document (Goddard 1976:87–88; Bowern 2008:185–188, 191–192).[3]

Crucially, the fieldworker will have to learn to read and interpret materials in very poor and pre-phonemic transcriptions, and to avoid undue normalization. Advice and examples regarding this are in Goddard (1976:76–78), Haas (1975), Austin and Crowley (1995), Voorhis (1996:462–469), and Bowern (2008:189–190). To give one example from another area of the world, in older documents in South Asian languages, it is common practice for orthographic *aa* to stand not for [a:] but for [a], and for *a* to stand for [ə].

The native language of the collector always needs to be taken into account. Somehow, the only two word lists of Apache languages ever collected on Mexican

[2] Hall (1969:321) appears to be the first to suggest that we need a practical manual for philology comparable to the fieldwork guides. We do not know of any recent general manual for the practice of philology. Therefore, manuals like Thieberger's for all parts of the world are a strong desideratum.

[3] This rule applies to the fieldworker's description and documentation, of course, but not to derived materials, such as pedagogical materials.

territory were written by Frenchmen (de Reuse 1994, 2006), and this, of course, has an influence on the way Apache is represented. The only record of Crimean Gothic was written by a Fleming, and the Flemish influence on his spellings of the Gothic words is obvious, and is critical to phonetic interpretation (Stearns 1978).

The philologist needs to be able to read many languages. Campbell (1997:8) points out, for example, that older sources in Algonquian or Iroquoian languages of the American Northeast can be written in Dutch, English, French, German, Latin, or Swedish. For the study of Tibetan documents, it would be ideal to know Chinese, English, Sanskrit, German, and Russian, and it will not hurt to know French, Hindi, Italian, and Hungarian either.

The unique characteristics of the contact language, particularly when this is not the native language of the collector, must also be taken into account. In John Gregory Bourke's Apache vocabulary, collected from Apache scouts and prisoners of war during his military campaigns against Apaches, there are a lot of Spanish words (Bourke 1980). These words can tell us something about the sort of Spanish the Apaches spoke with the U.S. military in the nineteenth century. The fieldworker John P. Harrington (see Chapter 3) used Spanish to elicit material from some of his Native Californian consultants, and his Spanish has been studied by Anderton (1991).

Even if the old records focus on an extinct language, it might still be useful for the fieldworker to know something about them, since that documentation can inform fieldwork on related languages, on languages originally in contact with the target language, or on languages typologically similar to it. For example, unusual typological characteristics have turned up in Troike's (1959, 1981, 1996) studies of Coahuilteco, an extinct language of South Texas. The linguistic value of closed corpora, i.e. relatively short bodies of data gathered from terminal speakers of languages now extinct, is discussed in detail by Grant (1991).

If the materials of the extinct language are very scant or poor, as is the case for parts of Australia and the New World, the best that can be done is to re-edit them and present them in a sketch form. This was done successfully for several languages of Australia, such as Woiwurrung, the language of the Melbourne area (Blake 1991), and for Tasmanian (Crowley and Dixon 1981). Grant (1994) presents all that is known of Karankawa, an extinct language of coastal Texas.

Occasionally, the philologist can encounter strange things in older documents or field notes. Campbell in his survey of the historical linguistics of the indigenous Americas (1997:13–14), describes all the known cases of fake or mistaken Native American languages. These include languages made up by linguists, languages made up by someone's consultant, or assumed names of languages that happen not to be names of languages at all.

5.2.2.2 Advice Regarding Historical Linguistics

If the fieldworker has historical interests, it is obviously necessary to compare the data found in old documents with fieldwork data for modern or related languages. Examples are the study of Vietnamese historical phonology by Gregerson (1981), the studies of Mayan historical phonology by Campbell (1978, 1990), work on Algonquian languages (Siebert 1975; Goddard 1976:77–81), and work on languages of New Caledonia and Polynesia by Hollyman (1974).

With due caution, interesting phonetic detail can be extracted from older materials. Canger (1990) is an example of philological phonetic work that can be done with the fairly numerous sixteenth century grammars of Nahuatl, some of which were quite perceptive (see also Section 3.2.2).

Sometimes, philologists do not pay sufficient attention to the detail of the early transcriptions. In the scant materials on Ofo, an extinct Siouan language of the American Southeast, aspiration of consonants was ignored by the first generation of historical linguists, but more careful scrutiny of Ofo materials shows that Ofo aspiration is genuine, and occurs where one expects it to historically (Haas 1975; Goddard 1976:77; de Reuse 1981; Rankin 2006).

Other historical studies focus on the grouping of languages into families. Among many studies based on philology, we cite as an example Güldemann (2006), a recent study of the San (formerly Bushman) languages of Southern Namibia.

5.2.2.3 Advice Regarding Editions and Studies of Older Texts

Editions of texts are probably the most fruitful way in which fieldworkers and philologists can collaborate. A good example, with detailed philological commentary, is Montler (1996), on Songish, a Salish language of British Columbia. Some of the most interesting old texts are unpublished catechisms, confessionals, and other religious materials written by Spanish speaking missionaries for many of the languages of Latin America. For many of the extinct languages of California, Baja California, and Southern Texas, these are the only text materials we will ever have. Good models for California are Beeler (1967) on Chumash and Beeler (1971) on Yokuts. For Baja California, we recommend Mixco (1978) on Cochimí, and Zamponi (2004) on Waikuri. For South Texas the work by Troike (1959, 1981, 1996) already mentioned in Section 5.2.2.1 is exemplary.

Editions of previously published manuscripts are also becoming more common in Latin America and the Caribbean. An example is Ruz and Birrichaga (1997), an edition of Zoque religious and grammatical materials from colonial Chiapas, Mexico. A superb example of an edition with full apparatus (commentary) is in van Rossem and van der Voort (1996), a collection of texts in Negerhollands, a Dutch-based creole formerly spoken on the Virgin Islands.

Editions of unpublished texts written by native speakers are of particular interest, due to their potentially high reliability. A good example is the *Native writings in Massachusett* (Algonquian), edited by Goddard and Bragdon (1988). Goddard's (2006, 2007) reeditions of Fox (Algonquian) texts written by native speakers, formerly published in a Boasian format, and informed by recent fieldwork, are also exemplary. DeMallie (1999) is an account of Lakota (Siouan) texts written by native speakers and edited by native speakers.

As far as studies of texts are concerned, philological approaches as applied to texts in traditionally unwritten languages were first described and labeled as “anthropological philology” by Hymes (1965) in a review of work on Chinook (American Northwest Coast) texts by the linguist, folklorist, and anthropologist Melville Jacobs. It is still valuable reading for the modern philologist.

Rudes (1999) emphasizes the importance of collaboration with native speakers when studying older texts in Tuscarora (Iroquoian), and Wolfart Christoph (1999) makes the same point regarding Cree (Algonquian) texts. Parks (1999) is a detailed study of the history of the study of Pawnee (Caddoan) texts. These three authors’ accounts are very much informed by their own field experience with these languages. See Section 13.4 for further discussion of the study of texts and text gathering.

5.2.2.4 Advice Regarding Lexical Materials and Dictionaries

It is also useful to study lexical materials in old records and to reorganize them in dictionary format, as has been done by for Mahican, an extinct Algonquian language of New York by Masthay (1991), for Kaskaskia-Illinois, another extinct Algonquian language by Masthay (2002), and for a variety of Tzotzil, a Mayan language of Mexico by Laughlin and Haviland (1988). Since Tzotzil is very much alive, the philological analysis has benefited, of course, from Laughlin’s and Haviland’s field experience. The Koyukon Athabascan language of Alaska is documented in a dictionary by Jetté and Jones (2000). Jetté was a Jesuit priest with a superb feel for ethnographic detail and a good understanding of the grammar; this was reorganized and added to by Jones, a native speaker linguist who also benefited from the experience of other Alaska Native Language Center fieldworkers.

It is also possible to publish manuscript vocabularies in chart form, following to some extent the original layout of the sources. An example is the Chumash and Costanoan vocabularies collected in California by the anthropologist H. W. Henshaw, and edited by Heizer (1955). In such a case, an alphabetical index of the words is a useful addition.

By lexicographical standards, the large majority of old dictionaries or lexical collections of endangered and unwritten languages tend to be rather short, so it is possible for a patient fieldworker (or better yet, a team of fieldworkers and consultants) to re-elicit or check on them within the lifetime of the consultants or while the language is still spoken. There are exceptions to this, however. The Yahgan language of Tierra del Fuego has a dictionary of 664 pages in two columns (Bridges 1933), but at the time of this writing there is only one elderly speaker of the language, so it is quite unfeasible to re-elicit more than selected portions of this dictionary. The South Andamanese Dictionary of Man (1923) has 136 pages in two columns, and 100 pages of appendices with words of great ethnographic and botanical interest. This variety of Andamanese is extinct; therefore, for Southern

Andamanese we have to rely entirely on philology. While Man (1923) might well be useful for further fieldwork on comparative Andamanese, it appears that there has not yet been any attempt to re-elicit words from the still-extant North Andamanese (Manoharan 1989:138–148, 168–175) on the basis of Man's masterful dictionary.

5.2.2.5 Advice Regarding Grammars

Another valuable endeavor is to pull together fieldwork notes or older grammars on a particular language, and to write a modern grammar on the basis of that material. This has been successfully done with two recently extinct languages of the Andes: Cholón of Peru (Alexander-Bakkerus 2005), and Uchumataqu of Bolivia (Hannß 2008). The benefit is that if these languages turn out not to be quite extinct, it will be easier to know what to look for in further fieldwork.

For North America, the voluminous field notes of John P. Harrington (see Section 3.1.3) will be the source of grammar writing for years to come. An example is Okrand (1977), a grammar of Mutsun (Costanoan of California), which is based on a portion of Harrington's field notes (about 2,500 pages out of a total of about 81,000 pages of Harrington's Mutsun notes). The 2,500 pages used by Okrand were primarily re-elicitations by Harrington of the phrasebook of an early nineteenth century Spanish missionary (Okrand 1977:7–8).

Another useful activity is the re-edition of older already published grammars, written in older European, Boasian, Sapirian, or Bloomfieldian traditions (Section 3.1.3). Examples are Sommerfelt's fieldwork-based grammar of a Breton dialect (Sommerfelt 1978), and the Micmac (Algonquian) grammar of Father Pacifique (Hewson and Francis 1990). Again, in these cases the fieldwork experience of the editors themselves aided in the re-editions, even though the original authors were already excellent fieldworkers. More of this work urgently needs to be carried out, as it contributes to a better understanding of past fieldwork techniques. One case of a published grammar that should be re-studied to extract material in an optimal way is Machoni (1877), a grammar of Lule, an extinct language of northern Argentina, which also contains religious texts without translation. A foretaste of the interesting typological features of this language is provided in Adelaar and Muysken's (2004:385–391) survey of Andean languages.

5.2.2.6 Advice on Re-eliciting on the Basis of Older Sources

The fieldworker should carefully prepare before going over old records with consultants. S/he needs to be cautious in case the materials deal with relatives or enemies of the speaker, or if sensitive subjects (such as religious topics, embarrassing gossip, or taboos on the deceased) are involved. Legal and ethical restrictions on how archival material may be used by the fieldworker is discussed at length in Bowern (2008:185–189).

Some consultants will prefer the set-up where the fieldworker reads material out loud to them, while others will prefer to look over the fieldworker's shoulder and try to decipher it together. One has to develop a feel for what an individual consultant is most comfortable with. One account of re-reading previously collected texts to consultants is in Voegelin (1954).

If the material one is reviewing is considered sacred or of high prestige, it will sound very presumptuous for the fieldworker to suggest redoing or rewriting the whole thing. But one can ask questions about specific portions chosen in advance. If the document merits respect, the consultant will of course appreciate the fieldworker's respectful treatment of it. On the other hand, early sources might be considered worthless or hardly legible by consultants, although considered precious by the fieldworker. The quality of nineteenth century field notes on Apache is so low that my [de Reuse's] consultants generally enjoyed tearing them apart.

Lessons from re-eliciting words from older sources in Tuscarora (Iroquoian) are discussed incisively by Rudes (2002:191–193). He notes that older vocabulary might have been replaced in the modern language, and he points out various possible collector-speaker miscommunications. For instance, the recorder might ask for a non-existent word; the speaker might misconstrue the question; the recorder miscopies the response, or the speaker might give a simplified "trade language" response. Rudes (2002:194–195) also provides a set of five rules for re-elicitation, given in Table 5.1.

Regarding step (5) in Table 5.1, it is important to note that Rudes' article was printed in a publication oriented towards the preservation of indigenous languages. So we interpret (5) as meaning, "omit from pedagogical or other materials intended for general consumption by the community". In fieldwork based on older records one should archive lists of words and forms either not understood or else rejected by one's consultant because these might be elucidated in the future. Even if it is

Table 5.1 Rudes' rules for reelicitation

1. Reelicit	Check older words with contemporary speakers if possible.[4]
2. Triangulate	Where a word is unknown in the modern language, look in other older sources to see if you find the word and confirm its prior existence.
3. Compare	Look at other, related languages and see if the word exists there.
4. Check credentials	If none of the three steps outlined above proves fruitful, it is still possible that the word in the older source is correct (...) In such cases it is necessary to examine the credentials of the researcher who collected the data.
5. Omit	In some cases, it may be necessary to omit questionable vocabulary from other sources from the dictionary. This decision should be made by contemporary speakers after all of the above efforts have failed.

[4] The material from this chart directly quotes Rudes (2002).

suspected that these forms were treated this way by the speaker because they may be sensitive in nature, it is still important to archive these forms somewhere, with the proper protections in place.

As Rudes (2002:194) also points out, when re-eliciting old words, caution is necessary when the consultant says an archaic form is impossible or ungrammatical because the pronunciation or meaning of the word might have changed, or the word may simply have fallen out of use. Consider that if one were to ask a linguistically-naive speaker of English to read and comment on handwritten English texts from the nineteenth century, quite a bit would be considered strange or barely comprehensible. Speakers of languages without a written tradition, not having seen an older version of their own language, might well have no concept that the language can change over a few centuries. So the older source and one's modern consultant could both be correct.

Conversely, if the sources to be re-elicited happen to be printed or written in a prestigious and authoritative looking sort of script, the power and prestige of the written word might have an effect on some speakers, and they might be more reluctant to say it is wrong. There might be an assumption that if the source is printed, it cannot or should not contain mistakes.

5.2.2.7 Advice Regarding Language Revitalization

From a community-support perspective, philological work on extinct or seriously endangered languages is also useful for modern descendants of the speakers who want to revive the language, or at least learn something about it. Of course, such work is not directly relevant to fieldwork methodology.

Discussions regarding the use of archival materials for Native American language revitalization are in Hinton (2001a, b), Rudes (2002), Bach (2004), and Warner et al. (2007). The most detailed set of instructions on using written records for revitalizating languages is in Nakayama (2007), providing as an example both published and unpublished materials by Edward Sapir on the Nuuchahnulth language (formerly Nootka, Wakashan family of the American Northwest Coast).

An intriguing case is the ongoing revival of Cornish, the Celtic language of Cornwall, which was extinct by the nineteenth century (Hinton 2001a:416). It is interesting to compare the first grammar of Cornish, based entirely on philological study (Jenner 1904), with the last (Wmffre 1998), based on a philological study of eighteenth century Cornish documents, the most recent on the language.

5.2.3 Conclusions

The last section of Goddard's (1976) foundational essay on philological approaches to Native American languages is called "Documentation", pointing out that the philologist also needs to look to the future. The article by Golla (1995), on the other hand, is about archiving and archived records, but also contains a helpful section on the need for interpreting the record with philological skills and historical knowledge (Golla 1995:152–154). Documentation is, of course, the creation of materials by modern fieldworkers, and it is good to remember that our modern archives may be the object of philological study by researchers hundreds of years from now. Such extension of a field philological approach to the future is also emphasized by Nakayama (2007:102–104).

The new view of language documentation, as reflected in Gippert et al. (2006) is geared entirely to the future, and – perhaps too optimistically – hardly discusses philology. Hill (2006), on the other hand, makes the often-forgotten point that the fieldworker is a second-language learner. As a second-language learner, the fieldworker may reach a point where s/he no longer needs to write glosses in his/her notes. Hill (2006:122) admiringly points out that in Ken Hale's (see also Section 3.5.4) field notes on Mountain Pima, (Uto-Aztecan of Northern Mexico), glosses were omitted after only six pages. The lesson to be drawn from this is that such practices make notes a lot harder for future fieldworker/philologists and for members of the target language community itself to interpret.

Philological work is, in a real sense, complementary to descriptive fieldwork, as already pointed out by Hall (1969). As more and more languages become endangered or extinct, language description will have to rely more on philology and – unfortunately – less on fieldwork. It is therefore important, as long as there is time to do both, that a fieldworker should also be a philologist, and that s/he should take the needs of future philologists into account.

5.3 Linguistic, Historical, Sociopolitical, and Cultural Preparation

Regardless of whether or not one intends to conduct an in-depth philological study of earlier materials, it is a good idea to read as much background as possible background on the target language. This includes grammars of related languages or of the language itself, as well as dictionaries, phrase books, text collections, and the like. To some linguists, this kind of preparation might make the field experience less interesting and might even be misleading. For instance, according to Fleck (2008:255–258), "Reading someone else's work before doing one's own research on the same language is like trying to finish a crossword puzzle that somebody else has started, with an undetermined number of mistakes in their answers". This may be a useful caution for beginning fieldworkers. A linguist with solid typological training however, should be able to judge the quality and veracity of the data and analysis found in earlier work. In addition, it would be a shame not to use these materials since they might be good sources of data; a lot can be learned by re-checking that data with consultants.

It also pays for the fieldworker to read as much as possible about the history and culture of the area s/he will be visiting. Especially useful may be dissertations by other fieldworkers who have worked in the same area (Feagin 2002:22). When doing fieldwork in linguistically diverse areas such as South Asia, it is necessary to know how the target language is affected by language contact (Gumperz 1964, 1972; Abbi 2001). It is also necessary to see if and to what extent speakers are bilingual or multilingual, since these factors will affect their fluency or their lexical choices in certain domains of interaction. The reasons for bilingualism - was it imposed or did it arise naturally because of historical circumstance - should be reviewed through reading about the linguistic and political history of the region, particularly as reflected in language policies. My [Chelliah's] Lamkang (Tibeto-Burman of India)-speaking consultants freely code-switch and code-mix with Manipuri, and if I were unable to tease apart the two languages, my resulting understanding of Lamkang would be inaccurate. This type of preparation is equally important for the Indian and non-Indian linguist, as many educated Indians are not aware of the details of the complex linguistic and social situations outside the major metropolitan areas.

An understanding of the political situation and the socio-economic and socio-structural differences between rural and urban areas would also help with understanding why language loss and language shift occur. In order to know which and why some varieties are stigmatized, it is necessary to know about national language policies and language status. In India, according to Abbi (2001:4–22), speakers of "scheduled languages"[5] often consider non-scheduled languages – even those spoken by more than five million speakers – to be inferior languages, spoken by less-advanced and thus inferior communities. To be sure, this stigma influences native speakers' attitudes towards their language. For example, speakers of a non-scheduled language may claim to be native speakers of the dominant language rather than the minority language.

To reinforce the importance of reading about a society before entering it, let us consider the importance of the caste system in daily interactions in India. This de jure outlawed but de facto observed method of social discrimination and separation may not be immediately observable to people from outside the community. Adherence to caste rules may also be stronger or more overt in rural areas. In any case, the fieldworker must know – for both rural and urban areas – where caste situates one's consultants in the social hierarchy. This provides crucial cues on: (1) how the fieldworker should behave with regard to food, seating, clothing, and the use of linguistic politeness indicators (2) how others in the community see the consultants' social standing and value the work done with those consultants, and (3) what caste-based dialect features exist in the variety being elicited. For these same reasons, it is also important to take the religious affiliation of a speaker into consideration. For example, in urban areas, Indian families that are religiously conservative tend to use a vernacular language in most interactive domains, whereas more westernized speakers – whether conservatively Christian or Westernized and secular – use English in many interactive domains. The modes of bilingualism used in the community must be understood before consultants are selected.

[5] Ray (2000:41) lists the following requirements for inclusion of a language in Schedule VIII. The language must: (a) be spoken by a significant number of speakers (where "significant" is defined as a majority of the population in a single geographical zone); (b) have an independent literary tradition with its own script (the interest in reviving *Meithei Mayek* and replacing the Bengali script is motivated by these requirements); (c) be recognized as the official language of an Indian state; and (d) be a "classical" language of India, i.e., be a language of Indian culture and heritage.

5.4 Preparing to Learn a New Language and Script

In most cases, fieldwork is conducted through a contact language and it is beneficial to gain some fluency in the contact language before going to the field. This could be accomplished relatively quickly by taking a formal course at a language center, or by participating in a study-abroad program, or through immersion. For administrative, personal, and fieldworker purposes, the usefulness of being proficient in a language used by the target language community is obvious. Even if a translator is used to mediate between the contact and target language, the fieldworker will need to be able to evaluate how well the translator is transmitting messages or representing consultants' responses. Many times, as pointed out in Newman and Ratliff (2001:5), so much effort is put into learning a contact language that the target language gets ignored. See also Burling ([1984] 2000:107–112) on learning contact languages in an urban setting.

Even more than learning the contact language, there are obvious advantages to learning the target language at the same time as making it an object of study. Everett (2001:170), who encourages the monolingual method of linguistic fieldwork, says, "… if one does not speak a language, one is working with a self-imposed handicap. Why should anyone want to turn down the clues, insights, intuitions, and constant grammar-learning and practice inherent in language-learning if one is genuinely concerned with a deep professional understanding of (aspects of) the language in question?" Other reasons for the fieldworker to try to learn the target language are that s/he can:

- Quickly increase the vocabulary and phrases s/he is exposed to
- Attempt monolingual elicitation thereby increasing the number of speakers contributing to the corpus
- Integrate with and demonstrate a commitment to the community (McLaughlin and Sall 2001; Moore 2009)
- Get grammatical information from real-life contexts (Moore 2009)
- Understand pragmatically motivated constructions (Gil 2001)
- Produce language forms to encourage speakers to "rethink" in their language when that language is moribund (Evans 2001)
- Gain respect rather than scorn from community members who are likely themselves multilingual (Moore 2009)
- Recognize and repair communicative breakdowns (Moore 2009)

Not everyone has the aptitude to learn a language, even given strong motivation to do so (Bley-Vroman 1990). Furthermore, some people are afraid of making mistakes and therefore do not progress rapidly as learners. Crowley (2007:157) points out that it is often easier to learn to speak by conversing with children, because with them the learner may not be afraid of making mistakes. Children may also correct the fieldworker with no embarrassment. Adults, on the other hand, are more likely to either tolerate mistakes or to poke fun at errors. The extent of natural input also determines success in language learning. Crowley (2007:157), for example, observes that he was able to attain fluency in Paamese (Austronesian of Vanuatu) because of two 6-month uninterrupted visits with the community, but he did not attain fluency in Erromango where he worked for three 2-month periods. Some other factors are: the size of the speech community; whether speakers are monolingual or also speak a lingua franca to which they switch into when talking to the field researcher; whether speakers take pride in their language or are embarrassed that they speak a non-prestige variety (Crowley 2007:158–159).

If the fieldworker wants to read earlier work on the target language or simply function in the speech community, it may well be necessary to learn a new script. In India, for example, there are 14 different scripts in use; to conduct fieldwork on an Indian language the fieldworker will probably have to learn one of them. In addition to reading previous linguistic work on the target language, access to written documents will help increase a corpus to include written samples such as newspapers, pamphlets, comic strips, and marriage announcements. In turn, written documents will lead to an understanding of the prescriptive standards observed in written genres. Also, as discussed by Schneider (2009), the fieldworker would need to know the language in order to prepare literacy materials.

A tip for speeding up learning the target or contact language is to begin with "chunk learning", whereby the learner can begin to communicate quickly by using formulaic expressions and other common collocations. It is useful to collect some common conversational scripts and 'small talk' lexica early on during fieldwork to help with fluency in conversations. Based on her experiences in northern Cameroon, Moore (2009) advocates "folktale socialization", where rote memorization of texts is used to increase fluency in the target language. Moore used texts "to train [her] mouth… learning bits of folktales and singing sounds embedded therein, and uttering proverbs that indexed folktales" (2009:248, 250). Moore also suggests reviewing video tapes of conversations in order to experience conversations at her own pace. A detailed discussion of how to learn a field language is Burling ([1984] 2000) where he provides these suggestions:

- Make communication primary in production and comprehension.
- Put yourself in a place where you can continuously hear the language.
- Memorize vocabulary efficiently.
- Involve consultants in language learning and teaching. Which words and constructions do they feel you should know?
- Isolate and spend time with "natural" language teachers.
- Practice by speaking to children.
- Record short passages from consultants and use these to build on vocabulary, intonation, and comprehension.
- Record conversations on known topics, and use these to build on vocabulary, intonation, and comprehension.
- Notice basic patterns such as the order of elements in the noun phrase, and practice these patterns in speech.

A note of warning: not all communities want researchers to gain communicative competence in their language. This is especially true in the North American contexts, where non-native speakers may be considered irritating wannabes, linguistic thieves, or show-offs. If command of the language is still rudimentary, speakers may cringe at hearing a broken form of their language. A note of encouragement: speakers will appreciate a fieldworker who can speak their language because, as Dobrin (2008:318) puts it, speaking the consultant's language indicates "that outsiders not only have things to say to them, but that they are also capable of listening to them."

5.5 Practical Preparation

The practical preparation before going to the field is similar to preparing for a study-abroad program.

5.5.1 Applying for Funding

The reality is that linguistic fieldwork takes money. However, it does not always require a large amount of funding, since the main expenses are equipment (which can often be borrowed from another researcher or departmental pool), travel (which, for a pilot study, could be a trip to a closer area or urban area), and payment to consultants (which can be negotiated so that the fieldworker can keep within budget). I [Chelliah] began with a student travel grant of less than $100, and a borrowed recorder and left over tapes from someone's fieldwork grant! If it is necessary to put together a team of researchers, however, substantial support will be required. Good sources to read on this topic are Abbi (2001), Ladefoged (2003), and Bowern (2008). A general resource for funding in linguistics is Peters (1986). Some tips on being successful with funding from our experiences are given below.

The fieldworker should identify appropriate funding agencies. For years, funding for linguistics projects was focused on research on generative grammar (see Newmeyer and Emonds (1971) for an interesting historical perspective on this topic). More recently, the scope of the type of work funded has broadened considerably.

At the time of this writing, there are several agencies that fund endangered language documentation.[6] Government and private agencies, such as the American Institute of Indian Studies, might fund linguistic descriptive fieldwork if the work is related to a community they are supporting. The appropriate program officers should be contacted to make sure that the proposed project is something their institution is interested in funding.

Deadlines should be checked. For most grants, applications usually have to be submitted 6 months to a year in advance. In addition, when applying through a host institution, it may be necessary to allow for an extra 2 weeks while the proposal

---

[6] See http://www.linguistics.ucsb.edu/faculty/infield/courses/resources/gw_funding_agencies.pdf for a useful list.

makes the rounds through various administrative offices. Since it typically takes granting agencies about a year to come back with a decision, there should be some contingency plan in place if completion of a thesis or tenure decision is based on the completion of fieldwork. It also helps to apply for multiple grants turn down one grant if another one comes through.

The fieldworker should read proposals which have been funded; these can be obtained from colleagues or advisors. Successful proposals are often given on funding institutions' websites and can be used as models. Reviewer comments on proposals are also useful, so we recommend finding find a fellow linguist who is willing to share a summary of these comments.

Here are some reasons why proposals are not funded: (1) the project is too ambitious and cannot be completed in the allotted time; (2) the principal investigator does not have a track record in terms of publications or other academic products; (3) the narrative reflects a poor understanding of the research situation or topic to be investigated; (4) the budget is not well justified; (5) the proper permissions have not been acquired; (6) other projects need urgent funding while this one can wait; (7) the timetable for various components of the project is not clearly delineated; (8) the role of the project participants is not clearly defined; (9) the proposal does not have community support; (10) the project is not ambitious enough or too much time/money is requested for what is going to be done; (11) the project does not plan to integrate philological work into the documentation project, i.e. there is no plan to study the old literature by early colonials or missionaries because they are hard to read, hard to access, or are written in a language not known to the principal investigator (see Section 5.2); (12) the proposed project is a duplication of effort, because another linguist is conducting a similar project; (13) the narrative displays a disagreeable or ad hominem attitude towards colleagues in related fields; (14) the proposal has a hidden agenda unrelated to scientific work, i.e. missionary work or pedagogical work; and/or (15) the proposed methodology or equipment do not represent best practice.

The fieldworker should obtain appropriate permissions or demonstrate feasibility for procuring permissions to visit the field site. It is advisable to conduct a pilot study or at least visit the community in which the study is to be conducted, so that the fieldworker can demonstrate community support for his/her project. These days, funding agencies usually require evidence that researchers have applied for Internal Review Board (IRB) clearance. It may also be necessary to submit letters of support from the institution that will host the grant, from a sponsoring institution in the field, and/or indication of permission to work with the community from appropriate community leaders. Evidence of government clearance or evidence of application for such clearance may be required by the funding agency. For more on IRBs and community permissions, see Section 5.8.4.

The main part of the proposal is a narrative which clearly sets out a plan of why, how, and when data is to be collected. This includes reference to previous related studies, methodologies to be applied, and the feasibility of the proposed work. The following is a list of questions that should be anticipated in a proposal for a fieldwork project. These questions are reworded versions of a list in Bowern (2008:171–172). We have also added a few questions of our own.

1. Information about the language and its speakers should be provided: Where is the language spoken? Which language family does it belong to? How many speakers are there? What evidence is there that the language is endangered?
2. Where will the project take place?
3. Has a pilot study been conducted? Has contact been made with the community? Has the applicant already published on this language or related languages?
4. What is the focus of the proposed study?
5. What will the products be: for example, a grammar, annotated archive of texts, phonetic study?
6. What methodology will be used to achieve project goals: for example, participant observation, interviews, participatory group research?
7. How will data be processed and analyzed? How much data will be collected (in terms of hours of recorded speech)?
8. What is the timetable and schedule of activities for the proposed project?
9. What benefits and consequences will the project and the project results have for the speech community and the academic community?
10. What supports the choice of a particular primary investigator over another: for example, does one researcher have previous work experience with the community, or know the contact language?

A well-justified budget should be constructed. Again, it is a good idea to read a sample budget. Putting together a budget is a time consuming process, so this should not be left for the last minute. Here are the major items to include:

- Travel to the site or sites: Some funding agencies may require the budget to reflect travel by the most inexpensive economy fare. The cost of airfare and other transport should be updated as close to grant submission time as possible since these can change quite drastically. Specify what travel expenses may be incurred within the site: bus, car, small plane, snowmachine, bikes or mules,[7] and car rental insurance.
- Lodging: It is often better to use an "actual cost" quote when calculating how much will be needed for lodging, but this will not always be possible. See Abbi (2001) for helpful suggestions on the types of places one can stay in India, also these are relevant for other parts of the world: local families, hostels, the guest-house of a corporation or think tank, YMCA or YWCA, government owned guest-houses, and school dormitories that may be vacant over vacations. Payment for lodging at these institutions may be fixed or negotiable. Local help may be needed in negotiating a fair price for lodging, and in the case of a host family this negotiating has to be done with tact. University per diem rates are typically higher than required and should generally not be used. When I [Chelliah] read proposals for work in India, and see budgeting for $150 per day for lodging, I am immediately wary of the accuracy of the rest of the budget; $150 is reasonable for 1 night at a four star

[7] Abbi (2001) says that bikes and mules are often the best method for travel in village India.

hotel but this is not the kind of establishment where a fieldworker is likely to stay.

- Living stipend: Although there may be a standard rate per diem allowed for expenses, the amount requested for daily living expenses should be reasonable.
- Payment for consultants' travel and time: this is a must if the consultant has to travel any distance to get to the fieldwork site. The consultant must be reimbursed for travel costs, time for travel, and time away from his or her regular job. See Section 7.12 for establishing how much to pay consultants and for various categories of personnel that might need to be hired.
- Communication expenses: It is useful to have a phone with an international calling plan that can function in remote areas. A disposable cell phone for field assistants could also be budgeted for. It will also be necessary to determine the availability and cost of Internet service, a useful tool for communicating with home base and for emailing copies of data files for safekeeping. Today, wireless broadband Internet access through USB modems makes Internet access easy. The hardware, monthly service charge, and start-up costs will need to be budgeted for.
- Photocopying or xeroxing, printing of field notes and/or purchase of relevant books.
- Equipment and supplies (see Section 5.5.2).
- Archiving expenses

In many cases, grant applications must be accompanied by letters of reference from an advisor or expert in the field. These should be requested well in advance. As a reference-letter writer, I need more than a week's notice to write a good letter; I need time to review relevant materials, ask the applicant questions, and make suggestions about revisions to the proposal.

The fieldworker should be persistent. Many grant proposals are funded on the second or third try. Each time a proposal is rejected, the applicant receives helpful referee notes that can be used to revise and resubmit the proposal. Another strategy is to apply to more than one granting agency, keeping in mind that it is usually not possible to hold two grants at the same time.

5.5.2 Things to Take

A fieldworker should start making decisions about the type of equipment to use for a field project well before the actual field trip. The equipment should be consistent with field methodology and goals. For example, if one plans to take a video recorder, how will the video recordings fit into plans for language description and analysis? Will software be used to align video with sound and transcription? If a team of field assistants is employed, more than one recorder and perhaps more than one computer will be necessary. In addition to equipment, which expendables and personal items will be required for the duration of the field trip? In many parts of the world, especially if one is not finicky about using a familiar brand, many items can be bought at the field site unless that field site is truly off the beaten path.

5.5.2.1 Equipment

The minimum equipment needed is two recorders (one of which can be a laptop), a good microphone, a backup microphone, and a notebook. However, there are three reasons why the field worker is expected to take more than the minimum amount of equipment.

First, as the goal of linguistic fieldwork often includes documentation of endangered languages, any recording that the fieldworker does must be of the highest quality, befitting the permanent and in some cases the only record of that language.

Second, laptops, top-of-the-line recording equipment, and supplies are readily available to fieldworkers sometimes within and certainly outside the developing world. Most funders support the purchase of such equipment, and this reflects an expectation that sophisticated equipment should be used in linguistic fieldwork.

Finally, current expectations are that publications that result from phonetic and discourse fieldwork are based on instrumental study (see Chapter 10), and must, as Ladefoged (2003) puts it, “be supported with sophisticated recordings, experimental design, and replicable and quantified analysis”. Today’s phonetic fieldworker, then, must learn about best practices in recording, analyzing, and storing sound data.

For specific recommendations on equipment and an equipment checklist, the reader is encouraged to refer to websites such as the one accompanying Bowern (2008),[8] the E-MELD website,[9] and Dwyer and Ono (2008), which includes specific recommendations on audio and video recording along with information on brands and models of recorders, recording media, microphones, cables, plugs, jacks, tripods, and hardware and software for video recording. It would also be useful to attend a preparatory workshop such as *InField*, the Institute of Field Linguistics and Language Documentation,[10] or the 3L International Summer School in Language Documentation and Description held in Leiden and hosted by the consortium of University of Lyon, Leiden University and the School of Oriental and African Studies, London; or workshops at the Language Documentation Training Center run by the Department of Linguistics, University of Hawai’i at Mānoa. In other words, there are several opportunities for training in this area before going to the field. See Bowern (2008:32–33) on the same point.

Since the specifics change constantly, we do not review brand or model recommendations here. Instead, we provide general considerations for selecting and using appropriate equipment for fieldwork. The fieldworker will want to take to the field a laptop or netbook with relevant software, at least one digital recorder, and at

---

[8] http://www.ruf.rice.edu/~bowern/fieldwork

[9] http://emeld.org/school/classroom/index.html

[10] The first InField was held at the University of California, Santa Barbara in 2008. The second was held at the University of Oregon in the summer of 2010.

least one high quality microphone. Consideration of the physical and social environment and the actual speech context further shape decisions on what equipment to take.

The Physical Environment

When conducting fieldwork in a developed country, the durability of equipment is not a major concern, since damaged equipment can be replaced relatively easily. On the other hand, if fieldwork is to be conducted in the rice fields of South-east Asia, or a rain forest in Brazil, equipment will be hard to replace and at the same time will undergo more wear and tear. We have heard some hair-raising stories about this; for example, Marc Brunelle (p.c.) tells us that his laptop was temporarily taken over by a colony of ants. Equipment needs to be evaluated in terms of how well it will stand up to the particular environmental challenges of the field site; backup equipment is needed in case primary equipment malfunctions, is damaged, or is stolen. On this point, Ladefoged (2003:189) recommends taking two of everything.

My consultants [Chelliah's] like to take control of the computer during elicitation sessions. I have found it best to give clear instructions on how a computer should be handled: don't tap on the screen, make sure your hands are dry, and tap – don't hit on – the keys or mouse pad. These simple instructions made it possible for us to share this precious equipment without worry on my part or embarrassment on theirs. It is important for the fieldworker to retain control of the recorder; there is no point in being shy about something basic to the success of the project. If there is a playback recorder used for transcription work, the consultant can take control of that once the original recording has been backed up for storage and archiving.

In addition to the challenges of dust, moisture, and heat, a recurrent problem in many parts of the world is the lack of or surge of electric current. In India, it seems prudent to work on battery at certain times of day when the voltage is high. Whenever possible, electrical equipment should be recharged only when connected to a heavy duty surge protector and voltage stabilizer. For this reason, when selecting a computer, extended battery life is a desirable feature. It is also useful to bring along more than one battery in case one gets damaged.

It is strongly recommended that the fieldworker bring rechargeable batteries and a recharger to the field since it is not always easy to find long lasting batteries. Another hard-to-find item, even in places like Fairbanks, Alaska, are older or the newest types of memory cards for recorders and digital cameras. It is a good idea to take extra memory cards.

While some things may be available, as a rule, it seems to be prudent to assume that no useful equipment is available at the field site. We strongly recommend discussing hardware and power requirements with others who have done linguistic or anthropological research in the same area as the planned fieldwork. This will prevent the waste of precious time trying to buy specific items and will avoid the disappointment when those items are not found.

The Social Environment

We consider the social environment as one of the most important factors in deciding what and how much equipment to take to the field. An urban or rural setting in a developed country that has a stable government, predictable execution of law and order, and minor disparity between rich and poor, gives the fieldworker the freedom to freely carry top-of-the-line equipment to the field. However, in many parts of the world, a fieldworker carrying expensive equipment becomes a prime target for theft and all the dangers that come with that crime. Additionally, expensive equipment sets the fieldworker apart from the community into which s/he is trying to fit. In these situations, the fieldworker will want to select equipment that is:

Inconspicuous and portable: Solid state flash drive recorders are light and don't look like they have a lot of bells and whistles. As long as the fieldworker resists the temptation to boast about the quality of the recorder, no one is going to know how valuable that piece of equipment is. A huge advantage of today's smaller recorders is that they are easy to carry around in a purse or backpack. Fieldworkers commonly use laptop computers to play back utterances for transcription work. We have heard of several instances where community members request the fieldworker to sell or give them the laptop. This would be good opportunity to think out loud about how a computer for the community rather than the individual might be an appropriate repayment for speakers' hospitality towards the fieldworker.

Replaceable: Again, for recorders and microphones it is advisable to carry replacements. In a pinch, lower quality equipment can be found in most places; analog recorders and analog tapes are almost universally available. However, since analog recorders should be avoided for work on endangered languages, taking a second digital recorder and microphone is advisable.

There is one use for less than top-of-the-line equipment: an analog recorder or cheaper digital recorder can be given away without making a big dent in the project budget. It is also useful to have recorders for use as community recorders so that members of the community can record samples of speech for the project in natural environments without the fieldworker being present. Even in these situations, the fieldworker should ideally provide field assistants with the best digital recording devices available. It is useful to keep in mind that the need for immediate cash often overrules the need for good recordings, and consultants may be tempted to sell the equipment rather than use it for the intended purpose. In this case, the fieldworker could double the chances of getting some decent recordings by providing the community with two \$300 digital recorders rather than one \$600 digital recorder.

The Speech Event

A final consideration when selecting equipment is the type of speech event to be recorded: What is the setting? How many speakers will be recorded and what type of interview will take place? What is the significance of the recording for analysis or the permanent record? To illustrate this, consider the following four scenarios for phonetic fieldwork:

- Scenario 1: A native speaker reads from a word list to be used for later acoustic analysis.
- Scenario 2: A fieldworker seeks native speaker input to confirm phonetic transcriptions and phonological hypotheses.
- Scenario 3: Narrative performances are recorded for advanced phonetic and phonological fieldwork.
- Scenario 4: A conversation is recorded for the study of fast speech phenomena.

The tables below provide equipment recommendations appropriate for each of these scenarios.

Phonetic Fieldwork Scenario 1: The native speaker reads a word list for later acoustic analysis	
Microphone[11]	• Channel (need a single channel microphone, stereo is acceptable). • Condenser type (need a clear signal transferred to the recorder). • Power (require long battery life). • Unidirectional (ensures that little background noise is recorded). • Placement (ensures recording of a steady sound with uniform intensity and no direct blasts of released air. A stand or head mounted microphone can be used depending on what the native speaker is comfortable with. The speaker should not be allowed to hold the microphone since it is usually not possible to hold the microphone at the same distance from the mouth. The microphone should not be too close to the recorder. • An internal microphone should only be used as a last resort. • Voice activation should be off. Any pauses in speech will stop the recorder, and there is danger that the first part of the next utterance will not be recorded.
Digital audio recorder	• Digital recording required, compact flash recorder preferred. • (1) The signal should not be compressed; record in WAV format with a sampling rate of 44,100 Hz and a bit depth of at least 16 bits (2) The recorder should not generate its own noise. Note that most computers do generate a high level of noise.
Backup audio recorder	There are many things that can go wrong when recording, and most often it is impossible to recreate what has been accomplished at a given session. A second digital recorder or a computer can be used for backup. These should be easy to operate.

(continued)

[11] See Nathan (2004) for a detailed discussion of how to select a microphone for fieldwork.

(continued)

Digital video recorder	Digital video recorder with external shotgun unidirectional microphone. A tripod should be used to set up the camera for optimal placement. Before the session begins, the camera should focus on the face and lips. The frame should allow for the recording of gesture and the placement of captions. The device should be operated by remote control so as to be less intrusive in the narrative event.
Computer	Laptop with a DVD burner and ample hard-drive space for backup. The appropriate software for analysis and database management should have been installed, tested and prepared before commencing fieldwork. The camera and recorder drivers and file transfer software should be installed and working correctly.

Phonetic Fieldwork Scenario 2: The fieldworker seeks native speaker input to confirm phonetic transcriptions and phonological hypotheses	
Microphone	Same requirements as for Scenario 1. Some additional comments about microphone placement: Since the recording will not be used primarily for acoustic analysis, an omnidirectional lapel microphone can be used. The speaker should be instructed on how to clip or place the microphone on the lapel themselves. If the consultant is not wearing clothing, the microphone can be suspended from a string looped around the speaker's neck. When working with several speakers a flat table microphone or several lapel microphones can be used. An adapter will be required to plug in more than one microphone into the recorder. Microphone batteries should be checked when they come with their own power supply. See Crowley (2007:123) for what happens when this is not done!
Digital audio recorder	The requirements for Scenario 1 hold if working on an endangered language and if the speaker tends to give a lot of supplemental information or produce narratives or songs during elicitation sessions.
Audio playback through laptop computer	The fieldworker plays back previously recorded utterances and records native speaker comments on the played back segments. It is preferable to play back from the computer and record with the digital recorder.
Speakers and headphones	It is useful to have both external speakers and/or headphones for sound playback and transcription work. Bowern recommends that headphones should have good frequency response but should not be noise cancelling. She also recommends that they not be worn for long periods of time due to possible ear infections in humid climates (2008: 23) An adapter will be needed if more than one headphone is to be plugged into a device.
Video recorder	The requirements for Scenario 1 hold. Sessions to be recorded should be carefully selected as video files take a lot of storage space. Catalog all video recordings carefully so that less relevant recordings can be set aside.

Phonetic Fieldwork Scenario 3: Narrative performances in natural settings for archiving	
Microphone, digital audio and video recorder	Same requirements as for Scenario 1. Equipment should be carefully selected and properly used as this data will be part of the record used for grammatical analysis, language documentation and revitalization, and language pedagogy.

Phonetic Fieldwork Scenario 4: A conversation is recorded for the study of fast speech phenomena	
Microphones, digital audio and video recorder	Same requirements as for Scenario 1. Microphone placement and unobtrusiveness of the recording equipment are of greatest importance here. The microphone should be placed so all speakers can be heard equally well or all speakers should be miked. The recording equipment should not make speakers self-conscious and hinder the free flow of conversation. See Section 10.3 for discussion.

To summarize, decisions on what equipment to take to the field should be based on a careful ethnography of the fieldwork situation, taking into consideration the physical and social environment; a review of the interview situation and the tasks that will be implemented; and the resulting data and its uses for the project, the speakers, and the language record to which the project will contribute.

5.5.2.2 Travel Checklist

When travelling internationally, we suggest starting with an existing checklist of necessary items for fieldwork, and tailoring that list to specific needs. Some useful checklists are: the Whole Earth Provisions backpacking list;[12] the checklist for anthropological fieldwork available on James Fox's website;[13] and Burgel R.M. Levy's fieldwork checklist.[14] See Bowern (2008) for needs specific to working in Australia, and Abbi (2001) for India. Both authors have useful suggestions relevant to all field situations. Here is a list of some commonly needed supplies:

Supplies for Data collection, Storage, and Dissemination:

1. Paper and notebooks: When my consultants [Chelliah] write in notebooks or on paper, they tend to fill every square inch of the page. There is no room left for translation or comments. This is because paper is a precious commodity to them; it is often unavailable because shops are closed due to strikes or other political

[12] http://idahoptv.org/outdoors/shows/next/pack.cfm

[13] http://www.stanford.edu/~popolvuh/field-checklist.htm

[14] http://www2.hawaii.edu/~faehndri/

unrest. To enable consultants to take useful field notes leaving wide margins and white space for comments and translations, it is a good idea to take paper or extra notebooks for their use. The ideal notebook for fieldwork should have lined pages and be spiral bound or bound so that the notebook can be written in whether or not a table is available, i.e. balanced on the knee or standing up (Dixon 2010:320).
2. Slate chalk-boards: Abbi (2001) says that these are useful in Indian rural settings so villagers can describe items in the material culture through drawings.
3. Postage: stamps, envelopes, cardboard boxes (often not locally available), sealing tape, scotch tape, glue, scissors.
4. Writing implements: pencils, sharpeners, erasers, pens, markers, highlighters, crayons, chalk.
5. Backup media: flash disks, external hard drives, compact disks (CDs) for data dissemination. Always carry some older form of backup, such as CD rather than a DVD since consultants may not have access to computers that can read newer media.
6. Batteries.

Personal Items:

1. Cooking implements: When cooking one's own food, it will be necessary to purchase a stove and some pots and pans. Take a fork, knife, and spoon; forks are not universally available. A Swiss army knife is always useful.
2. Emergency lighting: flashlights; hands-freeing head-lamps available at camping equipment stores; a penlight; candles and matches (they always work and don't require batteries!)
3. Medical supplies: first aid kit; anti-malarial drugs; anti-bacterial ointment; aspirin; diarrhea preventative; anti-itch ointment for insect bites; prescription medication (this should be planned for a few months in advance, to be prepared for an extended stay).
4. Mosquito nets and repellent (especially useful are the moist towels that can be carried in your pocket for emergency touch ups).
5. Clothing: lightweight clothing that is easy to wash and dry (dark clothing may dry faster than whites); at least one set of permanent press formal clothing; clothes that can be layered for warmth; for hot climates light cotton clothing that can cover the body (for cultures where shorts or short sleeves are unacceptable and to protect against sunburn); extra shoes and sandals, especially sandals reserved for indoor wear.
6. Toiletries
7. Bedding: Abbi (2001) notes that it is a good idea in most of India to take your own sheets, pillowcases and blankets to supplement what is provided in hostels and the like. Many families will not have extra bedding and may go without so that the fieldworker/guest doesn't need to. A down sleeping bag is a practical addition for cold climates.
8. Portable water filter and/or water purifying tablets.
9. Tent: Gerd Jendraschek notes that a tent can be set up inside a permanent structure to allow for some amount of privacy if needed. In addition, a tent can provide some

protection against mosquitoes and other bugs, leaking hut roofs, can be locked to keep belongings secure (Research Center for Linguistic Typology 2009:21).

Communication and Networking

1. Cell phones: find out what identification is necessary to buy cell phone service that will function at the field site. In India, you will need several extra passport photographs to get phone service. It may be necessary to buy a satellite phone (Bowern 2008). Consider buying a phone with a limited number of minutes for your main consultant; this may be the only sure way of keeping in touch and getting started with a field session in a predictable fashion.
2. Business cards.
3. Maps (topographical, political, linguistic, road maps).
4. Gifts: It is a good idea to bring an assortment of lightweight gifts appropriate for males and females of several different age groups. Someone who has been to the area could provide appropriate advice on what gifts to bring. Linguistics books and articles are appreciated by local linguists.

In addition to *what* to take, the fieldworker must consider *how much* to take, and *when* to take it. Should some things be shipped in advance, and only items crucial to data collection be carried. Another issue is how much to take. For a first time visit to a particular field site, it is best to take as much as possible because it is never certain what is going to be available unless one has been able to check with local residents or other researchers. On my [Chelliah's] first fieldtrip to Manipur, I took all that I needed for fieldwork with me – recorder, notebooks and the like. I shipped my clothing and other personal items separately. Unfortunately, that shipment took 2 weeks to arrive. Fortunately, everything I needed was available in town. Unfortunately there was a general curfew in place, and I was unable to go to the market to buy basic necessities. A useful lesson learned, and a mistake never repeated.

5.5.2.3 Travel Documents

A valid passport and visa are required for travel to most international venues. Plenty of time should be allowed for travel documents to be processed. The passport should be valid at least for the duration of the fieldtrip but preferably for up to 6 months after travel to the field site. It should not be assumed that a visa is not required: even if a tourist visa is not needed, a research visa or some documentation showing clearance for work in the country may be necessary. It might be useful to have an official letter from an institutional head explaining the purpose of the fieldtrip and expressing support for the proposed project to show immigration officials. As a rule, it is a bad idea to visit a country under false pretenses; not only is it illegal but also, if the fieldworker is found out, s/he may have trouble with subsequent visits and may make things difficult for future researchers.

Check travel warnings on your government's website to make sure that there are no restrictions for travel to the area you want to visit, and be aware that different countries do not issue identical travel advisories. Funding or institutional support may be withheld because the area where the proposed research is to take place is considered unstable. We recently heard of a proposal to conduct linguistic fieldwork in a volatile part of India that was quashed within the researcher's university when an administrator looked up the region on the U.S. Department of State website on international travel.

It may also be necessary to show proof of a health exam or specific test such as for HIV or tuberculosis. All the required immunizations should be updated and recorded in an official document such as an International Certificate of Vaccination. For those who haven't traveled much, check recommended immunizations with the relevant home country's state department. Useful sites for this information are the Center for Disease Control at http://www.cdc.gov/, the World Health Organization at http://www/who/int/en/, and the Travel Doctor at http://www.tmvc.com.au/.

As with all travel, tickets should be purchased well in advance, and the cost of extra luggage should be estimated and included in budgeting.

5.6 Psychological Preparation

It is good to be forewarned that the field experience will not be an easy experience. We refer the reader to Newman and Ratliff (2001:8), who list the following factors that could make the fieldwork experience difficult:

1. Lack of linguistic preparation
2. Trouble learning the target or contact languages
3. Breakdown of equipment
4. Disease and other health issues
5. Food and housing issues
6. Money: the social distance it creates between the fieldworker and consultants, and obligations to use money for social services
7. Incompetent or dishonest consultants
8. Problems adapting to or agreeing with local culture
9. Dealing with the death of speakers, especially when there are few speakers remaining
10. Problems with local government officials, or with other linguists, anthropologists, aid workers, or missionaries in the area
11. Worry about family at home or family brought along to the field site
12. Boredom and loneliness, lack of personal space and lack of privacy

The articles in Newman and Ratliff (2001) consider these issues and other relevant articles are referenced there as well. See also Crowley (2007:161–165), who after 3 months in the field began to feel the lack of privacy, fear the slow progress on his dissertation work, and feel a desire for the familiar. Feelings of culture shock will pass, and may be appeased by a quick vacation away from the field site. The precise problems faced by a fieldworker will only become clear in the field, but it is a good idea to talk to other linguists who have worked in a similar situation to get a heads up on what to expect. If one is going to do field work in Thailand, one cannot base one's idea of what it is going to be like in the field by talking to someone who does fieldwork in Australia. Many of the issues will be the same, but many will be different. For instance, challenges relating to physical comfort, disease, food, or equipment breakdowns will be relevant in a developing country, but not in the same way or to the same extent as in Western Europe, where equipment can be replaced with a phone call or Internet order, and food and other physical concerns are easily taken care of. For practical advice on what to expect regarding health, accommodation, and food, see Bowern (2008: 139). For a personal account of these issues in West Africa, see Mc Laughlin and Sall (2001), and for Mexico, see Macaulay (2004).

Some fieldworkers feel that it is crucial to be in the field, surrounded by a community of speakers in their "natural" setting, to really understand and describe the language under study, engaging in what Dixon calls "immersion fieldwork" (Dixon 2007). Although it can be a mistake to try to become "one of the community," (Dixon 2010:317), some amount of socialization is necessary to correctly interpret utterances in their social context. A nice example of this is Riley (2009), where she finds that daily interactions in the language clarify for her the underlying intentions behind utterances and thus impact her data and analysis. Aikhenvald (2007:4) expresses this opinion well, saying that "…recording, learning, and analyzing a new language as it is spoken in its own environment is the most intellectually exciting and invigorating enterprise. Despite all the physical difficulties, frustrations and sometimes even dangers that living in an unknown environment may bring, you live through a whirlpool of discoveries and sudden flashes of understanding. This is what "safe and comfortable existence" within the "secure confines" of a familiar environment (as Burridge 2007… puts it) can scarcely provide."

An opposing view is voiced by Munro (2003:140) who states that "fieldwork can be done anywhere". She says that, "For the most part, linguistic data gathered away from speakers' traditional homelands can be just as valid as linguistic data gathered in those homelands."

So in which situations should the fieldworker feel justified in excusing themselves from the field? Consider the cases of serious health problems or danger due to political strife. In this case, the "interview" method of fieldwork – working with speakers outside the actual field site in an urban center, for instance may be a reasonable stop-gap measure. There are some obvious downsides to the interview method: the language cannot be observed as spoken by a wide variety of speakers; influences of a second language may produce a contact version of the target language; and the cultural richness of the language will be muted due to the absence of cultural events.

But giving up on or postponing fieldwork in an area where speakers are ready to participate in a documentation project is unnecessary. This would be succumbing to the notion that there is only one right way to conduct fieldwork. Here is a case in point: I [Chelliah] have been unable to plan an extended stay in Manipur state, India, for several years now, due to frequent kidnappings and a lack of law and order. In the midst of this, I was approached by a Manipuri linguist and later by interested speakers of Lamkang and asked to begin a Lamkang documentation and description project. We were able to obtain funding for the project, and it seemed that the time was ripe for this work. In alignment were the availability of funding, a Manipuri linguist's long term interest and availability, and the interest of respected speakers. Had we not taken advantage of the convergence of these factors, the opportunity may have passed forever. Thus, even though I could not be a participant-observer, our team was able to come up with a plan for tapping into the language resources of the community in Manipur (two project members are now trained to record and get translations for Lamkang texts), and for transcribing and analyzing the data with the help of a linguistic expert (a team meets outside of Manipur regularly to process the data collected). This is not an ideal situation, but it is at least, a start.

As observed by Hildebrandt (2007), phonetic fieldwork may benefit from being carried out off-site. She says, of her work in Nepal on Manange, "Most of my work in Manang is in villages along the shores of the (roaring) Marsyangdi River or in villages on the slopes of the Annapurna mountain range, where windy afternoons are common. The outdoor areas that are sheltered from the wind and water typically contain other ambient noise-producing elements like domestic animals and children. Working indoors brings its own challenges, as houses in upper Manang consist of one or two-room dwellings, built mainly of stone, a naturally reverberating material. With these environmental challenges, one solution is to apply a more liberal definition of "fieldwork" for phonetic purposes. I personally have found that better results are obtained when language informants perform phonetic elicitation work away from daily existence contexts."

Examples of successful fieldtrips where linguists did not live with a speech community are related in Dixon (1984) and for Ken Hale in O'Grady (2001:227) by whom brief surveys of several languages were done in addition to deeper studies of individual languages. No doubt, the trips where Hale is said to have recorded data in five languages in 3 days also count as fieldwork. Of course, there are too many instances of where a speech community does not exist. If there are only a few speakers remaining and there is no day-to-day communication in the target language, it might be advantageous to live off-site where an office can be set up for data analysis (Crowley 2007:152).

An example of difficulties with consultants was related to us by Sadaf Munshi. In her work with Burushaski speakers in Kashmir, she feels the pressure to conform to conservative dress codes, requiring covering the head as well as other expected modes of conduct for women (p.c.). Munshi's position within this society is both "insider and outsider" (Dwyer 2006:36). She is an insider because as a native Kashmiri she is a member of the larger community of Srinagar where Burushaski is spoken. Furthermore, as a Muslim, she shares religious traditions and practices with the Burushaski community. But she is simultaneously situated outside the community because she is not a Burushaski. An equally strong factor giving her an "outsider" status are her long held ideological beliefs that women should not be required to cover their heads. This is a practice that she has actively disavowed within her own community where such pressures also exist. However, in order to work within the Burushaski community, she must suppress her outsider self – which holds one set of beliefs dearly and has fought for and succeeded in attaining them in her personal life – and privilege her insider status. A fieldworker must be prepared to juggle two ways of behaving which may mean they cannot always feel they are being true to themselves.

A fieldworker cannot expect to be liked by everyone in the field. Bowern points out that some people may not like fieldworkers because of previous negative experiences with researchers who have taken advantage of them (Bowern 2008:164). In addition, other researchers in the area may be jealous of or wary of a newcomer's successes, or they may be curious about or unimpressed by the newcomer's integration into native culture.

The lives of one's consultants may not be easy. War, internal conflicts and terrorism; diseases such as malaria, tuberculosis, and AIDS; hunger; drug and alcohol abuse (often leading to crime, gang warfare, and car accidents), are all conditions that a fieldworker may encounter in the field. Fieldworkers must be prepared to experience these conditions through their consultant's lives. Linguists are often left with the question of how they can help their consultants. It is not always possible to change lives simply by dint of wanting to do so. Bowern (2008:165) asks: "Would your plans for improving your consultant's lives actually solve any of their 'problems'? Would a half-attempted solution in fact make things worse, both for the community and for you?".

5.7 Preparing for Emergencies

As discussed in some detail in Research Center for Linguistic Typology (2009:10), fieldworkers must learn how to manage risk and deal with incidents such as civil unrest, fire or theft, injury, personal attack, physical or mental health issues, and adverse weather conditions. A little planning and forethought in this area will be of great help as one or more of the challenges posed by these issues may well arise during a fieldtrip.

1. Identify and cultivate a support system: It is important to identify one or two trustworthy people at the field site. If possible, it can be helpful to befriend a clan leader, village chief, or official who is known to and answers to someone outside the community. Others who may be of help are local missionaries. In many parts of the world, the police are corrupt and feel that they are above the law, so it should not be assumed that the police would be the first place to turn to in case of an emergency.
2. Plan for emergency health care: Locals contacts or other researchers in the area should be contacted about health care arrangements before the need for medical care arises. The Research Center of Linguistic Typology manual (2009:6) also advises fieldworkers to take a first-aid course in injuries incurred in the wilderness, or to read about village health care, as in Werner (1977). Since prescription drugs such as antibiotics may be adulterated or past the expiration date, the fieldworker should purchase medications from a trustworthy source.
3. Keep back-ups of travel and insurance documents: It is always possible that a passport, visa, research permit, ticket, or important contact information is lost or stolen. If copies of these documents are deposited with an advisor, peer,

or someone trusted at the field site, information from the documents can be retrieved easily to help with the replacement process.

4. Observe and follow local customs with regard to displaying and storing personal possessions: It is customary in many parts of the world to lock valuables away when not in use. It is best to follow this practice and to be very liberal in what one defines as "valuable", as can be seen by this report from Yvonne Treis: "In Ethiopia, I learned that people made a difference between "taking" and "stealing". If an item that is not locked away goes missing, it is not considered a serious matter. In fact the owner is considered to be at fault for not taking appropriate care. It is only considered 'true' stealing if an item that is locked away in a box, cupboard, or suitcase is taken." (Research Center for Linguistic Typology 2009:12).
5. Observe, read, and follow local customs: Many misunderstandings and potentially dangerous situations can be avoided by following local norms of behavior in terms of dress, personal space, eye contact, physical contact, smoking, drug and alcohol use, and sleeping conditions (e.g. should women live and sleep alone, or should they house with a host family?).
6. Ensure a steady source of income: Some thought should be given to how cash can be carried and stored. Can your home institution wire-transfer money in the local currency to a major city center to which you have relatively easy access? Even when seemingly foolproof arrangements are made something can go wrong. Sara Hale, Ken Hale's wife, recalls about the Hale's first few months in Australia, "…we weren't getting our money from NSF, but finally we heard that the secretary at the Anthropology Department in Sydney where our cheques were being sent was in the process of having a nervous breakdown, and she'd been throwing all the mail into the rubbish. We didn't get a penny from NSF for six months" (Hale 2001:20).
7. Listen to local advice: If there is civil unrest or dangerous weather conditions, the local population will know best how to handle the situation. Similarly, it is best to ask for and heed advice about places or people to be avoided.

5.8 Permissions

Crowley (2007:70) says, "A linguist should (…) not just turn up on the beach or at the airstrip and expect immediately to move into data gathering mode." There are several layers of permission that must be secured before fieldwork can begin, and not receiving even one of those could throw a monkey wrench in the entire fieldwork project. Project feasibility is therefore closely linked to the feasibility and success of acquiring permissions.

5.8.1 Central and Local Governments

The first step is to check whether the country to be visited issues research visas or requires research permits. The fieldworker should check with other linguists to find out how difficult it is for researchers to enter the country. How long does it take to get a reply or have an application processed? We sometimes think that if we fill out the required forms and honestly and earnestly detail what we are going to do and how our study will benefit the communities we work with, then permissions will be granted simply on the merit of our case. But this is not how the world works. The motive of a government official in granting or denying a request for a research visa may be completely unrelated to the merit of the research (Foley 2002:132). The reasons may have to do with the bureaucrat's past experience with other researchers; internal departmental politics; the political or social situation of the community of study; a prelude to a bribe request; lost paperwork; or just the slowness of bureaucracy. Given that permissions are not always granted in a timely fashion and sometimes not at all, what should the fieldworker do?

First and foremost, as already mentioned in this chapter, the fieldworker should not consider working without permission. Without it, the proposed research and the research of following linguists will suffer (Dimmendaal 2001). Instead, the fieldworker should investigate legal avenues for greasing the wheels of bureaucracy. For example, a local university or other research institution might be willing to advocate on behalf of the fieldworker, or might be willing to contact somebody who knows somebody who can put in a good word for the project. The fieldworker should be prepared to switch to "plan B" if things do not fall into place. For example, if permission to work on a particular language in India is not forthcoming, perhaps the same language or a closely related language could be studied as spoken in Bangladesh. If permissions to work in a rural area are not immediately granted, another option is to begin fieldwork with speakers in an urban setting as a stopgap measure. It will still be necessary, even in this circumstance, to get permissions from speakers, and through them from their community, to work on the language. So this measure is of little help if the community is opposed to the fieldworker working on their language. Standards were more lax in the past. When Wick Miller saw that it was impossible for him (Miller) to carry out fieldwork on the Acoma (Keresan of New Mexico) language, he worked with a speaker in the Los Angeles area. While Miller's work on Acoma is unsurpassed, he is criticized by modern Keresan language activists for having done this (Sims and Valiquette 2000).

The next step is getting permission from a local authority to conduct linguistic fieldwork. It is possible to get a green light from the central government but to be stopped by a local authority. There may be many reasons for this; for example, the local government might assume that the fieldworker is spying for the central government (Barnes 1963); local officials might see the fieldworker as benefiting the 'troublemakers' in the community, for instance, a minority group that is vying for an independent government or more autonomy. Here, careful explanation of exactly what the fieldworker intends to do and not do must be repeated. For example, if there are ongoing arguments over water or land rights, the fieldworker may need to clarify that s/he will not be taking sides or documenting histories that will support one claim over the other. Speaking from a South Asian perspective, it is best to have a third party negotiate these permissions for the fieldworker. An educational or research institution is often the best connection for this.

5.8.2 Tribal or Cultural Councils

Local community-level permissions, can, in fact, be the trickiest of all to negotiate. The actual shape of the ruling body could be a cultural council, elected tribal council or simply a group of elders. The obvious method of securing permission is to identify the center of power in the community – say a tribal council – and have the proposed linguistic fieldwork project added to the agenda item of a regularly scheduled council meeting. One might expect permission to be readily granted after the council has considered exactly what the project data will be used for, and who will be given credit, and has been assured that speakers will receive compensation for time and effort; and that all resulting materials will be made available to tribe. Unfortunately, there are often complications. In many cases, the procedure in place to deal with such requests may be vague and only inconsistently enforced. There may also be disagreement within the community about whether permissions should ever be granted. In the southwestern United States, for example, Native American tribes may reject a request because it may seem to them that speakers will receive little or no benefit from the work, while researchers are imagined to make a great deal of money from resulting books and products. After all, a vibrant tourist trade in world-famous Southwestern jewelry, pottery, and rugs has made many non-Indian dealers very wealthy, so there is an understandable concern that linguists might profit in a similar measure from Native languages as a uniquely Southwestern cultural resource.

It can be difficult to find out who exactly is in power. It may be a village chief or tribal elder, but sometimes the center of power may be hidden. As Paul (1953:431) states, "In some groups the men who mold opinions are not always tagged for all to see." A possible pitfall in this is that the fieldworker might approach less powerful people for permission, and thereby inadvertently snub those in power. Alternatively, another researcher could have the ear of the council, and could act as a gatekeeper to the community. The petitioning researcher might need to get permission from different people who are antagonistic towards each other, or from different villages where one agrees with project goals and plans but the other does not. Factions may exist along religious lines – more conservative against more westernized – or may divide those who feel that the language should be written down from those who do not. It is up to the fieldworker to find out where community members or blocs stand on these issues.

What, then, is the best way to secure permissions from a local ruling body? It is important to tread softly and take some time in establishing relationships with community members before going to the council with a request for permission. Paul (1953:431) describes the fieldworker's role-setting period as a game, saying, "His is the strategy of a player…He cannot predict the precise plays which the other side will make, but he anticipates them as best he can and makes his moves accordingly." The moves of the game are introductions and individual approvals which then allow the fieldworker to move closer to the center of power. Paul (1953:430) describes it in this way:

> There is no prescription for finding the correct entrée into a new community. It depends on the sophistication of the community and the amount of advance information the investigator is able to get. Frequently he can count on a chain of introductions which

> leads at least to the threshold of his group. By the time he reaches a provincial center or trading post near his destination, he is likely to have learned the names of people who have contacts with the natives. Here on the peripheries he can pick up bits of information which will serve to orient him. The novice who is anxious to obtain the full acceptance of the natives sometimes by-passes regional administrators for fear of prejudicing his reception. But it will do him little good to be well received by the natives, only to be impeded by higher authorities who make it their business to follow the movements of strangers. Apprised of his plans, power figures can come to his aid in case of unexpected trouble; uninformed, they can cause trouble of mistaken conceptions as to his intentions. Whether he asks them to provide an introduction will depend on this estimation of how such auspices will promote or prejudice his standing in the community to be studied.

Seeking permissions to work on Native American tribal reservations in the southwestern United States similarly requires an introductory period. One approach is to meet and get to know Native Americans who are interested in their own language (but who are not necessarily native speakers) on some neutral ground, such as a language preservation conference. Here, the fieldworker can get to know students/teachers/educators, and also the older speakers or elders who are invited to give presentations at such conferences: s/he can talk to as many people as possible to gauge who s/he can develop a connection with and can try to find out speakers' opinions about language preservation work. After meeting with people off the reservation, the fieldworker could visit the same students/teachers/educators on the reservation, and then hopefully be introduced to other tribal educators or schoolteachers. It is important to realize that an introduction is just an introduction, no matter how warm and friendly. It does not count as having gained extra support in the community. It is necessary to give people on the reservation plenty of time to look the fieldworker over for themselves, interact with him/her, tease him/her (often mercilessly), and make sure that s/he is okay, before a sort of relationship/friendship can be developed.

Outsiders will often tell the fieldworker: "Go see so and so, s/he is a nice guy/gal, you will get along". That approach does not work on the reservation; people decide for themselves who they consider a nice guy/gal. Every outsider on Southwestern reservations, no matter his/her references, goes through a period of testing and teasing to see what they are about and to see if they can keep their cool even under embarrassing circumstances. It is also good to meet tribal administrators, although these tend to be more suspicious about what one wants, and will initially have little time to listen to the fieldworker.

After making some friends and supporters among the educators on the reservation, the next step is to inquire about the permissions needed. The requests should be kept simple when presented to the tribal council. Supporters should be in attendance.

The fieldworker should not be a linguistic show-off; it might seem like a great idea to sprinkle one's speech with target language forms that one has learned through books or from friends, but council members might react negatively since, especially when a language is endangered and not heard on a regular basis, it may sound strange to have these almost-sacred sounding words coming out of an outsider's mouth.

Be aware of who is on the tribal council. If the majority on the tribal council is friends with a faction that does not like the people the fieldworker associates with, it is better to wait until the composition of the tribal council changes. Being friends with the tribal chairman/president/governor helps, but s/he by him/herself cannot grant the permission. In order for the permission to be valid, it needs to be voted on by the tribal council. To summarize, it takes time to acquire permissions at the local level. The fieldworker must get to know speakers and let the community get to know him or her. If one tribal council is unfavourable, a later council may not be, so useful attributes for this requirement are persistence and patience.

If the fieldworker does not get permission at the local level, is it possible to simply move his/her project off site? The short answer is "no". S/he must get permission at the local level even if working with speakers outside the community. It is the ethical thing to do. Also, if the fieldworker does not seek permission, consultants may face community censure. We know of a speaker of a Native American language who agreed to be a consultant in a Field Methods class out of her home state. When word got around that she had "shared the language" without permission, she was criticized rather harshly by some of her own family members. In this community, as in others, "linguistic and cultural knowledge is viewed as corporate, so that individual community members are not in a position to consent to share materials with outsiders." (Linguistic Society of America Ethics Committee 2006). Furthermore, without permissions, a researcher could be prevented from publishing data, as a result descriptive and documentation efforts might not be available to the speech community or the academic community.

5.8.3 Gatekeepers

Once a community gives a fieldworker the go-ahead to begin fieldwork, appropriate speakers need to be found for the project. Even with all the required permissions, access to speakers may be guarded by gatekeepers – those who control, for better or worse, the participation of a speaker in research. This situation is especially possible in the case of an endangered language where there are only few older speakers available. Oftentimes, the gatekeeper is a younger relative or the elder's caretaker. These gatekeepers may or may not be justified in their concerns. It is possible that earlier researchers have behaved unethically towards the speaker by, for example, not paying them, tiring them out, or not acknowledging their contributions to a project. Or the gatekeeper may simply not want to share a precious resource that is under his or her care, the reasoning being that control of knowledge is control of power.

It may even be the case that the speakers themselves are enthusiastic about participating in a linguistic project, but a gatekeeper stands in their way. This situation can cause an ethical dilemma for the fieldworker. A speaker might want to work with the fieldworker, think that s/he will do useful work for the community, tell the fieldworker that the gatekeeper, in his/her hatred for linguists, is unreasonable in not giving permission, and encourage the fieldworker to ignore the gatekeeper and work with him/her behind the gatekeeper's back. What is a fieldworker to do in such a situation, especially if the language is terminally endangered and if it is unlikely that someone else will have the speaker's enthusiastic support? There is no simple solution to such a dilemma. The individual desires of the speakers, particularly if they feel they represent the community wishes better than the gatekeeper does, should be taken into account. However, if the fieldworker gives in to the desires of the speakers, s/he might cause unpleasantness for the speakers by taking sides in what is perceived as an intra-community conflict or even a family feud, and thereby expose themselves to legal reprisals from the gatekeeper.

Sadaf Munshi (p.c.) tells us of a different take on a similar gatekeeper situation in her fieldwork. Her own extremely kind and helpful consultant was a middle-aged male school teacher. His opinion, mirroring that of the larger society, was that women could not speak their native language as well as men simply because they were women. So Munshi found herself banned from working with women until she made friends with her "teacher's" daughters.

5.8.4 Home Institutions and Internal Review Boards

In the recent past, it was possible to conduct federally funded research in the United States without an ethical review from either the funding agency or the fund-administering academic institution. But today, in the United States it is necessary for the field linguist to obtain clearance from an Internal Review Board (IRB). According to Bowern (2008:148), ethics boards are also in place in Canada, Australia, the United Kingdom, Scandinavian countries, other European countries, Israel, and Japan. IRB reviews are required for research involving human subjects. Even though there are no invasive procedures involved for most linguistics projects, linguists must complete the IRB process since they work with human subjects. The purpose of the IRB is to protect the subjects of research projects conducted under the purview of the IRBs home institution.

There are usually two parts to an IRB application: a proposal to a funding agency and a blank informed consent form written specifically for that proposal. The consent form is to be read by each human subject that the linguist works with. It must be signed and in some cases separate sections are to be initialed. Here are the main points covered in a consent form:

1. **Title of Study: ___________________**
2. **Principal Investigator and affiliation:**
3. **Purpose of the Study:** (Tells the subject about the research project in which he or she is being asked to participate in non academic language.)
4. **Study Procedures:** (Explains exactly what the subject will be asked to do and the amount of time it will take for the subject to complete the task.)
5. **Foreseeable Risks:** (Explains the potential risks or discomfort involved in the study.)

6. **Benefits to the Subjects or Others:** (States what the benefits are to the subject, the speech community, Linguistics, or others who will benefit.)
7. **Compensation for Participants:** (Explains payment, when it will be received, in what form, and how much.)
8. **Procedures for Maintaining Confidentiality of Research Records:** (Describes the methods used to protect subject confidentiality/anonymity, and to secure storage of recordings and explains where the data will be archived and who will have access to the data.)
9. **Researcher Contact Information:** (Provided for use by subjects who might have questions about the study.)
10. **Statement of Review for the Protection of Participants:** (States that the proposal and consent form have been reviewed by an IRB.)
11. **Statement of Research Participants' Rights:** (Records a signature from the participating subject showing that they have understood the consent form and are aware of their rights.)
12. **Certification of Principal Investigator:** (Certifies that the consent form has been reviewed with the subject.)

It is easiest to complete IRB paperwork once the proposal (needed for answering (3)) and budget (needed for answering (7)) are written. Some reasons why IRB clearance is delayed or rejected are: the application is not complete; the language used in the narrative is too specialized for subjects to understand; payment to subjects is not satisfactorily explained; or subjects apparently cannot speak English but no provision has been made for translation of the consent form. Crowley (2007:25) advises that, because other academicians typically do not know anything about linguists, it may be difficult to convince them that, "the person who is going to tell you that the word for 'dog' in the Eastern Rainforest language is *siliwan* is not subjecting himself or herself to unspeakable potential harm." It is best to make the significance of interactions with speakers crystal clear to reviewers. Without approval from an IRB, funds will not be released to the researcher. It is also important to note that the idea of potential harm can be interpreted differently by researcher and speaker.

Most linguistic fieldworkers will have no trouble explaining why the requirement for written consent possess an unnecessary obstruction to linguistic fieldwork: for example, subjects may be illiterate; there may be a tradition of fearing signed contracts; mention of payment may make the transaction too formal and may imply obligations that the speaker is unwilling to commit to; the fieldworker may have only 30 min to record a speaker and there is not enough time to explain and have the form signed; and the relationship between the speaker and researcher is ongoing with conversations and meetings taking place over a period of time, so that it is hard to fix an exact amount for payment. Technically speaking, one consent form is necessary per person, per project, per year. Again, it may seem strange to a speaker from a non-industrialized community to have to sign the same paperwork repeatedly. Possible modifications of this process, making it more suitable for fieldwork with some speakers, are presented below.

In order to construct an informed consent form to fit the diverse needs of linguistic fieldwork, an IRB may be willing to allow for verbal rather than written consent. Verbal consent is useful for those unexpected situations when working with a speaker for a very short time; perhaps they are willing to provide a narrative or a record a conversation but will not be involved further in the project. In this case, the fieldworker could start recording but preface the recording with a question such as: *Today is DATE. I'm happy to be here with (prompt the speaker to say his or her name). NAME OF SPEAKER has agreed to tell us the story of how she moved to this area. This story will be used for our research and we will also make a copy for NAME OF SPEAKER and her family*. See Crowley (2007:29) for a similar script. Dwyer (2006:44) makes the interesting point that, to some, a verbal consent may actually feel more binding than written consent, since the required eye contact and physical co-presence suggest a personal responsibility on the part of the researcher to have the conditions of the agreement upheld. Dwyer (2006:44) continues: "…this is why many people (e.g. indigenous peoples of the Americas) find oral contracts more binding than written ones; written ones can be torn up and forgotten, but not ones sealed by physical contact."

In order to get informed consent, a field linguist must explain to speakers what the purpose of data collection is. Crowley (2007:27) advises against explaining data collection in linguistic terms, since speakers will not understand the terms used. Instead, he says, linguists should explain their "intentions in terms that people *are* likely to understand," saying, "you will…be attempting to 'learn the language by collecting words and stories'". There is a grey area here. While it is true that speakers will not understand the details of linguistic analysis and there is no point in confusing the issue by throwing around linguistic terminology, a simplistic representation of the goals of linguistic fieldwork may appear to purposely "hide" all the products that will follow from linguistic fieldwork. Some of these will have significant consequences for the community; for example, if the language does not have a writing system, the community might be encouraged to create one based on a phrase-book or grammar produced by the linguist. One approach is to discuss the various products of fieldwork in a formal setting before a tribal council or cultural committee, and once having their go-ahead, provide a simplified explanation to the individual contributor.

Not all speakers are satisfied with a simple explanation. In my [Chelliah's] experience, the Lamkang speakers I have worked with are very curious about my interest in Lamkang. In the typical scenario, I arrive at a speaker's home with my Lamkang field assistant and a Manipuri linguist colleague. I sit quietly in a corner, as does the Manipuri linguist. My Lamkang field assistant begins explaining what the project is about. The main point made is that, while school textbooks for Lamkang can be created by translating from Manipuri, Bengali, or English, the one thing that cannot be gotten through translation is a Lamkang grammar. Therefore an expert is needed to study the language and write materials on the basis of which a grammar can be written. The Lamkang speakers listen and ask questions. They want to know if we have permission from the community, and my Lamkang field assistant explains from which villages we have permission. Then the questions move to the involvement of outsiders: who is the Manipuri linguist? Why is he interested in us? Now, the language used switches to Manipuri, the lingua franca of Manipur. The Manipuri linguist explains how his father and a neighboring Lamkang village chief have been lifelong friends, how his interest in the Lamkang language has been encouraged for years by this village chief, and how he wants to continue to document traditional stories and songs at the behest of this village chief. Now, finally, it is my turn. Who is she? How does she come into the picture? Again, it is the Lamkang field assistant and the Manipuri linguist who explain how I am to help with Lamkang documentation. The entire process takes about 40 min. The actual narratives we record also take about 40 min. But this is time well spent because the next day we are given the names of two more speakers we can contact and work with. The process of explaining what we were doing was only slightly shorter this second time around, but different, in that the new speakers had already heard about what we were doing and encouraged us to continue because they understood the urgency of documenting the language.

Dwyer (2006:44–45) includes third-party consent as yet another type of acceptable consent. In this case, an intermediary might give consent for another person to be involved in a research project. This might be a gatekeeper (see Section 6.3.3), parent (see Bowern 2008:153–154), or a village chief.

When preparing a collection of texts or other cultural materials for wide dissemination, it is a good idea to "renew" permissions, if at all possible, through a quick email or call to the narrator/orator to remind them of their contribution (Crowley 2007:26–27).

As many of us have experienced, speakers can seemingly arbitrarily change their minds on what is acceptable. However, these changes-of-heart may in fact be related to speakers' growing understanding of what the project entails and how their data is to be used. That is, consent is not necessarily fully informed from the first explanation of the project. Sadaf Munshi (p.c.) tells us of an instance where she was given permission by some young women to videotape them conversing and to use the resulting data in her work. After the recording, however, the participants asked her not to use the data because they were more "relaxed" in their dress and demeanor than was socially acceptable. On two other occasions, speakers who were very hard to locate and record were finally recorded narrating some traditional stories. These speakers asked to listen to their narrations. Once they heard themselves, they asked that Munshi not use these narrations. Instead, they would provide her with better narratives the next time they met. That 'next time' never happened!

Dorian (2009) reflects that informed consent can never be fully informed from the speakers' perspective. Not only do speakers typically not understand the finer points of grammar that are of interest to linguists, but they may also not know how the data they share will impact local relationships. Fieldworkers themselves sometimes cannot understand local sensibilities even after many months of fieldwork. The fieldworker must realize that speakers' understanding of the project grows and changes and, therefore, a fieldworker must be ready, for example, to return tapes (as Dorian has had to do) or retract aspects of language of descriptions (as we have been requested to do with our genetic classification of a language spoken in

Northeast India). In our view, the fieldworker must use common sense in dealing with community concerns by carefully steering clear of community members who capriciously or deliberately (for any number of political, personal, and financial reasons) change their views on what is acceptable.

Some institutions require, as part of the IRB approval, that researchers guarantee that they will destroy all primary data at the end of a project end so as to ensure the privacy of subjects. This might be a good rule for some types of research, but it is obviously counterproductive for linguistic documentation and description. Therefore, the researcher must explain that the ethical move for descriptive linguistic fieldwork is archiving primary data in a repository and making sure that the data is not destroyed. See Crowley (2007:55) on this point.

5.9 Conclusion

As with any academic endeavor, the success of original research depends on preparatory reading and planning. Fieldwork is different from other academic work because it requires a major investment of self, and also involves human interaction and input on a highly personal matter – that is, consultants' language. Therefore, preparation for the field involves the practical (which can be managed with checklists) and the psychological (which can only be managed with honest reflection), and, since fieldwork crosses local, national, and political boundaries, it also requires negotiating with entities at these levels. All of this planning does not occur with only the fieldworker's goals in mind. The goals and desires of the communities and speakers where the target language is spoken must also figure in to the planning and execution of the research project. We cover this topic in Chapter 6.

References

Abbi, Anvita. 2001. *A Manual of Linguistic Field Work and Indian Language Structures.* (Lincom Handbooks in Linguistics 17.) Munich: Lincom Europa.

Adelaar, Willem F. H., with the collaboration of Pieter C. Muysken. 2004. *The Languages of the Andes*. (Cambridge Language Surveys.) Cambridge: Cambridge University Press.

Aikhenvald, Alexandra Y. 2007. Linguistic fieldwork: Setting the scene. *Sprachtypologie und Universalienforschung* (Focus on Linguistic Fieldwork, ed. by Alexandra Y. Aikhenvald), 60(1):3–11.

Alexander-Bakkerus, Astrid. 2005. *Eighteenth-Century Cholón*. Utrecht: Landelijke Onderzoekschool Taalwetenschap (LOT).

Anderton, Alice J. 1991. The Spanish of John P. Harrington's Kitanemuk Notes. *Anthropological Linguistics* 33(4):448–457.

Anttila, Raimo. 1979. Philology and Metascience. In *Linguistic Method. Essays in Honor of Herbert Penzel*, ed. by Irmengard Rauch and Gerald F. Carr, 497–507. The Hague: Mouton.

Austin, Peter K., and Terry Crowley. 1995. Interpreting old spelling. *Paper and Talk: A manual for Reconstituting materials in Australian Indigenous Languages from Historical Sources*, ed. by Nick Thieberger, 53–101. Canberra: Aboriginal Studies Press.

Bach, Emmon. 2004. Philology and endangered languages: The case of Western Abenaki. In *SOAS Working Papers,* ed. by Kirsty Rowan and Alex Bellem, 13:343–349.

Barnes, J. A. 1963. Some ethical problems in modern fieldwork. *British Journal of Sociology* 14:118–134.

Beeler, Madison S. 1967. *The Ventureño Confesionario of José Señán, O.F.M*. (University of California Publications in Linguistics 47.) Berkeley, CA and Los Angeles, CA: University of California Press.

Beeler, Madison S. 1971. Noptinte Yokuts. In *Studies in American Indian Languages*, ed. by Jesse Sawyer, 11–76. (University of California Publications in Linguistics 65.) Berkeley, CA and Los Angeles, CA: University of California Press.

Blake, Berry J. 1991. Woiwurrung, the Melbourne Language. In *Handbook of Australian Languages*, Vol. 4, ed. by R.M.W. Dixon and Barry J. Blake, 30-122. Oxford: Oxford University Press.

Bley-Vroman, Robert. 1990. The logical problem of foreign language learning. *Linguistic Analysis* 20(1–2):3–49.

Bourke, John Gregory. 1980. *Vocabulary of the Apache or 'Inde language of Arizona & New Mexico, collected by John Gregory Bourke in the 1870s and 1880s*, ed. by Carole J. Condie. Occasional Publications in Anthropology. Linguistics Series 7. Greeley, CO.: Museum of Anthropology, University of Northern Colorado.

Bowern, Claire. 2008. *Linguistic Fieldwork. A Practical Guide*. New York: Palgrave MacMillan.

Bowman, Elizabeth. 1959. An attempt at an analysis of Modern Yucatec from a small corpus of recorded speech. *Anthropological Linguistics* 1(4):43–86.

Bridges, Thomas. 1933. *Yamana-English. A Dictionary of the Speech of Tierra del Fuego*, ed. by Ferdinand Hestermann and Martin Gusinde. Mödling, Austria: Missionsdruckerei St. Gabriel.

Burling, Robbins. 2000. *Learning a Field Language*. 2nd edition Prospect Heights, IL: Waveland Press.

Burridge, Kate. 2007. A separate and peculiar people – fieldwork and the Pennsylvania Germans. *Sprachtypologie und Universalienforschung* (Focus on Linguistic Fieldwork, ed. by Alexandra Aikhenvald), 60(1):32–41.

Campbell, Lyle. 1978. Quichean Linguistics and Philology. In *Approaches to Language. Anthropological Issues*, ed. by William C. McCormack and Stephen A. Wurm, 223–233. The Hague and Paris: Mouton.

Campbell, Lyle. 1990. Philological Studies and Mayan languages. In *Historical Linguistics and Philology*, ed. by Jacek Fisiak, 87–105. Berlin and New York: Mouton de Gruyter.

Campbell, Lyle. 1997. *American Indian Languages. The Historical Linguistics of Native America*. New York and Oxford: Oxford University Press.

Canger, Una. 1990. Philology in America: Nahuatl: What loanwords and the early descriptions of Nahuatl show about stress, vowel length, and glottal stop in sixteenth century Nahuatl and Spanish. In *Historical Linguistics and Philology*, ed. by Jacek Fisiak, 107–118. Berlin and New York: Mouton de Gruyter.

Crowley, Terry. 2007. *Field Linguistics. A Beginner's Guide*. Oxford: Oxford University Press.

Crowley, Terry, and R. M. W. Dixon. 1981. Tasmanian. In *Handbook of Australian Languages*, Vol. 2, ed. by R. M. W. Dixon and Barry J. Blake, 395–421. Canberra: Australian National University Press.

DeMallie, Raymond J. 1999. 'George Sword Wrote These': Lakota Culture as Lakota Text. In *Theorizing the Americanist Tradition*, ed. by Lisa Philips Valentine and Regna Darnell, 245–258. Toronto: University of Toronto Press.

Dimmendaal, Gerrit J. 2001. Places and People: Fieldsites and informants. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 55–75. Cambridge: Cambridge University Press.

Dixon, R. M. W. 1984. *Searching for Aboriginal Languages. Memoirs of a Field Worker*. Chicago, IL: University of Chicago Press.

Dixon, R. M. W. 2007. Field Linguistics: A minor manual. *Sprachtypologie und Universalienforschung* (Focus on Linguistic Fieldwork, ed. by Alexandra Y. Aikhenvald) 60(1):12–31.

Dixon, R. M. W. 2010. *Basic Linguistic Theory. Volume 1 Methodology*. Oxford: Oxford University Press.

Dobrin, Lise M. 2008. Discussion Note: From linguistic elicitation to eliciting the linguist: Lessons in community empowerment from Melanesia. *Language* 84(2):300–324.

Dorian, Nancy C. 2009. Documentation and responsibility. *Language and Communication* 29(3):210–229. Online: http://top25.sciencedirect.com/subject/arts-and-humanities/2/journal/language-communication/02715309/archive/27/.

Dwyer, Arienne M. 2006. Ethics and Practicalities of cooperative fieldwork and analysis. In *Essentials of Language Documentation*, ed. by Jost Gippert, Nikolaus Himmelmann and Ulrike Mosel, 31–66. Berlin and New York: Mouton de Gruyter.

Dwyer, Arienne and Yoshi Ono. 2008. Audio recording equipment survey. Online: http://www.linguistics.ucsb.edu/faculty/infield/courses/resources/Infield08_DwyerOno_%20AudioEquipmentHandout.pdf.

Ehlich, Konrad. 1981. The native speaker's heritage: on the philology of "dead" languages. In *A Festschrift for Native Speaker*, ed. by, Florian Coulmas, 153–165. The Hague, Paris and New York: Mouton.

Evans, Nicholas. 2001. The last speaker is dead – long live the last speaker. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 250–281. Cambridge: Cambridge University Press.

Everett, Daniel L. 2001. Monolingual field research. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 166–188. Cambridge: Cambridge University Press.

Feagin, Crawford. 2002. Entering the Community: Fieldwork. In *The Handbook of Language Variation and Change*, ed. by Jack K. Chambers, Peter Trudgill, and Natalie Schilling-Estes, 20–39. Oxford: Blackwell.

Fleck, David W. 2008. Sugerencias methodológicas para realizar trabajo de campo lingüístico en la Amazonía. *Lexis* 32(2):251–280.

Foley, William A. 2002. Field methods. In *The Linguistics Encyclopedia*, Second edition, ed. by Kirsten Malmkjær, 131–137. London and New York: Routledge.

Gil, David. 2001. Escaping Eurocentrism: Fieldwork as a process of unlearning. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 102–132. Cambridge: Cambridge University Press.

Gippert, Jost, Nikolaus P. Himmelmann, and Ulrike Mosel, eds. 2006. *Essentials of Language Documentation*. (Trends in Linguistics. Studies and Monographs 178.) Berlin and New York: Mouton de Gruyter.

Goddard, R. Ives, and Kathleen J. Bragdon. 1988. *Native Writings in Massachusett*. 2 Vols. Philadelphia: The American Philosophical Society.

Goddard, R. Ives. 1976. Philological approaches to the study of North American Indian languages: Documents and documentation. In *Native Languages of the Americas*, Vol. 1, ed. by Thomas A. Sebeok, 73–91. New York: Plenum Press.

Goddard, R. Ives. 2006. *The Autobiography of a Meskwaki Woman. A New Edition and Translation*. Memoir 18. Winnipeg, Manitoba: Algonquian and Iroquoian Linguistics.

Goddard, R. Ives. 2007. *The Owl Sacred Pack. A New Edition and Translation of the Meskwaki manuscript of Alfred Kiyana*. Memoir 19. Winnipeg, Manitoba: Algonquian and Iroquoian Linguistics.

Golla, Victor. 1995. The Records of American Indian Linguistics. In *Preserving the anthropological record*, ed. by Sydel Silverman and Nancy Parezo J., 143–157. New York: Wenner-Gren Foundation for Anthropological Research, Inc.

Grant, Anthony P. 1991. Closed Corpora and "Circumstantial Evidence". In *Papers from the American Indian Languages Conferences Held at the University of California, Santa Cruz, July and August 1991*, ed. by James E. Redden, 37–66. (Occasional Papers in Linguistics 16.) Carbondale, IL.: Department of Linguistics, Southern Illinois University.

Grant, Anthony P. 1994. Karankawa linguistic materials. *Kansas Working Papers in Linguistics* 19(2):1–56.

Gregerson, Kenneth L. 1981. *A Study of Middle Vietnamese Phonology*. Dallas, TX: Summer Institute of Linguistics. [first edition 1969]

Güldemann, Tom. 2006. The San Languages of Southern Namibia: Linguistic appraisal with Special Reference to J. G. Krönlein's N|uusaa Data. *Anthropological Linguistics* 48(4): 369–95.

Gumperz, John J. 1964. Linguistic and Social Interaction in Two Communities. In *The Ethnography of Communicat*ion. *American Anthropologist* (Special Issue), ed. by John J. Gumperz and Dell Hymes, 66(6), II, 137–153. Menasha, WI: American Anthropological Association.

Gumperz, John J. 1972. Linguistic diversity in South Asia, Introduction (with Charles Ferguson). *International Journal of American Linguistics* 26:3, part 3, vii-18. (Reprinted in Charles Ferguson, ed., 1972. *Language Structure and Language Use*. Stanford, CA: Stanford University Press)

Haas, Mary R. 1975. Problems of American Indian Philology. In *Language and Texts*: *The Nature of Linguistic Evidence*, ed. by Herbert H. Paper, 89–106. Ann Arbor, MI: Center for Coordination of Ancient and Modern Studies, The University of Michigan.

Hale, Sara Whitaker. 2001. Reminiscences of the trip to Australia 1959–1961. In *Forty Years on: Ken Hale and Australian Languages*, ed. by Jane Simpson, David Nash, Mary Laughren, Peter Austin and Barry Alpher, 19–28. Canberra: Pacific Linguistics.

Hall, Robert A. Jr. 1969. Review of Samarin, William J., Field linguistics. A guide to linguistic field work. *Romance Philology* 22(3):317–321.

Hannß, Katja. 2008. *Uchumataqu. The lost language of the Urus of Bolivia. A grammatical description of the language as documented between 1894 and 1952.* Leiden: CNWS Publications.

Hanzeli, Victor Egon. 1969. *Missionary Linguistics in New France: A Study of Seventeenth- and Eighteenth-Century Descriptions of American Indian Languages*. The Hague: Mouton.

Heizer, R. F. 1955. *California Indian Linguistic Records. The Mission Indian Vocabularies of H. W. Henshaw*, ed., with ethnographic notes, by R. F. Heizer. Anthropological records 15:2. Berkeley, CA and Los Angeles, CA: University of California Press.

Hewson, John, and Bernard Francis (translators and transcribers). 1990. *The Micmac Grammar of Father Pacifique*. Memoir 7. Winnipeg, Manitoba: Algonquian and Iroquoian Linguistics.

Hildebrandt, Kristine A. 2007. Phonology and fieldwork in Nepal: problems and potentials. *Proceedings of Conference on Language Documentation and Linguistic Theory*, December 8–11, 2007, SOAS, ed. by Peter K. Austin, Oliver Bond, and David Nathan, 133–144. London: SOAS. Online: http://www.hrelp.org/publications/ldlt/papers/ldltproceedings.html

Hill, Jane H. 2006. The ethnography of language and language documentation. In *Essentials of Language Documentation*, ed. by Jost Gippert, Nikolaus Himmelmann and Ulrike Mosel, 113–128. Berlin and New York: Mouton de Gruyter.

Hinton, Leanne. 2001a. Sleeping Languages. Can they be awakened? In *The Green Book of Language Revitalization in Practice*. ed. by Leanne Hinton and Kenneth Hale, 413–417. San Diego, CA: Elsevier Science.

Hinton, Leanne. 2001b. The use of linguistic archives in language revitalization The native California language restoration workshop. In *The Green Book of Language Revitalization in Practice*. ed. by Leanne Hinton and Kenneth Hale, 419–423. San Diego, CA: Elsevier Science.

Hollyman, K. J. 1974. De L'Importance Et De La Valeur Des Premiers Documents Européens: Enquêtes en Nouvelle-Calédonie et aux Loyalty. In *Les langues sans tradition écrite; méthodes d'enquête et de description*. (Actes du Colloque International du CNRS, Nice 28 Juin-2 Juillet 1971.), 109–174. Paris: Société d'Études Linguistiques et Anthropologiques de France. No. 3, Numéro spécial.

Hymes, Dell H. 1965. The methods and tasks of anthropological philology illustrated with Clackamas Chinook). *Romance Philology* 19(2):325–340.

Jenner, Henry. 1904. *A Handbook of the Cornish Language Chiefly in Its Latest Stages with Some Account of its History and Literature*. London: David Nutt.

Jetté, Jules, and Eliza Jones. 2000. *Koyukon Athabaskan Dictionary*. Fairbanks, AK: Alaska Native Language Center, University of Alaska.

Karttunen, Frances. 1983. *An Analytical Dictionary of Nahuatl*. Austin, TX: University of Texas Press.

Krauss, Michael E. 2006. A history of Eyak language documentation and study: Fredericae de laguna in memoriam. *Arctic Anthropology* 43(2):172–218.

Ladefoged, Peter. 2003. *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*. Malden, MA: Blackwell.

Laughlin, Robert M., and John B. Haviland. 1988. *The Great Tzotzil Dictionary of Santo Domingo Zinacantán, with Grammatical Analysis and Historical Commentary*. Three vols. (Smithsonian Contributions to Anthropology Nr. 31.) Washington D.C.: Smithsonian Institution Press.

Levitt, M. L. 1989. Linguistics as history: Preserving linguistic oral records. *International Journal of American Linguistics* 55(4):417–423.

Linguistic Society of America Ethics Committee. 2006. *The LSA Ethics statement*. Online: http://lsaethics.wordpress.com/2008/07/

Macaulay, Monica. 2004. Training linguistics students for the realities of fieldwork. *Anthropological Linguistics* 46(2):194–209.

Machoni de Cerdeña, Antonio. 1877. *Arte y vocabulario de la lengua lule y tonocoté*. Buenos Aires: Pablo E. Coni.

Man, Edward H. 1923. *A Dictionary of the South Andaman (Âkà-Bêa) Language*. Bombay: British India Press.

Manoharan, S. 1989. *A Descriptive and Comparative Study of Andamanese language*. Calcutta: Anthropological Survey of India.

Masthay, Carl, ed. 1991. *Schmick's Mahican Dictionary*, with a Mahican historical phonology by David H. Pentland. (Memoirs of the American Philosophical Society, vol. 197.) Philadelphia: American Philosophical Society.

Masthay, Carl, ed. 2002. *Kaskaskia Illinois-to-French Dictionary*. Saint Louis, MO: Privately published.

Mc Laughlin, Fiona, and Thierno Seydou Sall. 2001. The give and take of fieldwork: noun classes and other concerns in Fatick, Senegal. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 189–210. Cambridge: Cambridge University Press.

Mixco, Mauricio J. 1978. *Cochimí and Proto-Yuman: Lexical and Syntactic Evidence for a New Language Family in Lower California*. (University of Utah Anthropological Papers 101.) Salt Lake City, UT: University of Utah Press.

Montler, Timothy. 1996. A reconstruction of the earliest songish text. *Anthropological Linguistics* 38(3):405–438.

Moore, Leslie C. 2009. On communicative competence...in the field. *Language and Communication* 29:244–253.

Munro, Pamela. 2003. Field linguistics. In *The Handbook of Linguistics*, ed. by Mark Aronoff and Janie Rees-Miller, 130–149. Malden, MA: Blackwell Publishing.

Nakayama, Toshihide. 2007. Using written records to revitalize North American languages. In *The Vanishing Languages of the Pacific Rim*, ed. by Osahito Miyaoka, Osamu Sakiyama, and Michael Krauss, 91–106. Oxford: Oxford University Press.

Nathan, David. 2004. Sound recording: microphones. *Language Archives Newsletter* 1(3):6–9. Online: http://www.mpi.nl/LAN/

Newman, Paul, and Martha Ratliff. 2001. Introduction. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 1–14. Cambridge: Cambridge University Press.

Newmeyer, Frederick J, and Joseph Emonds. 1971. The linguist in American society. *Chicago Linguistics Society* 7:285–305.

O'Grady. Geoff. 2001. Tribute to Ken Hale: our 1960 collaboration. In *Forty Years on: Ken Hale and Australian Languages*, ed. by Jane Simpson, David Nash, Mary Laughren, Peter Austin and Barry Alpher, 227–230. Canberra: Pacific Linguistics.

Okrand, Marc. 1977. Mutsun Grammar. Unpublished Ph.D. dissertation, University of California, Los Angeles, CA.

Parks, Douglas R. 1999. George A. Dorsey, James R. Murie, and the textual Documentation of Skiri Pawnee. In *Theorizing the Americanist Tradition*, ed. by Lisa Philips Valentine and Regna Darnell, 227–244. Toronto: University of Toronto Press.

Paul, Benjamin D. 1953. Interview techniques and field relationships. In *Anthropology Today: Selections*, ed. by Sol Tax. Chicago, IL: University of Chicago Press.

Peters, Ann M. 1986. *Handbook for Grant Proposal Preparation*. Washington, D.C.: Linguistic Society of America.

Rankin, Robert L. 2006. The interplay of synchronic and diachronic discovery in Siouan grammar-writing. In *Catching language: The Standing Challenge of Grammar Writing*. (Trends in Linguistics. Studies and Monographs 167.) ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 527–547. Berlin: Mouton de Gruyter.

Ray, Sohini. 2000. The Sacred Alphabet and the Divine Body: The Case of Meitei Mayek in North-Eastern India. Ph.D. dissertation, University of California, Los Angeles, CA.

Research Center for Linguistic Typology. 2009. *Fieldwork Manual. Fieldwork and Your Wellbeing*. Bundoora: La Trobe University, Research Center for Linguistic Typology.

de Reuse, Willem J. 1981. Grassmann's law in Ofo. *International Journal of American Linguistics* 47(3):243–244.

de Reuse, Willem J. 1994. Guillemin's "Mexican Apache" vocabulary. In *Memorias del II Encuentro de Lingüística en el Noroeste*, Vol II, 81–100. Hermosillo, Sonora, Mexico: Universidad de Sonora.

de Reuse, Willem J. 2006. Berlandier's Lipan Apache Vocabulary from Mexican Texas (1828). Paper read at the XI Encuentro de Lingüística en el Noroeste, Universidad de Sonora, Hermosillo, Mexico, November 2006.

Riley, Kathleen C. 2009. Who made the soup? Socializing the researcher and shaping her data. *Language and Communication* 29:254–270.

van Rossem, Cephas, and Hein van der Voort, eds. 1996. *Die Creol Taal. 250 Years of Negerhollands Texts*. Amsterdam, The Netherlands: Amsterdam University Press.

Rudes, Blair A. 1999. Why Collect Texts? The Native, Evangelical, and Americanist Traditions among the Tuscaroras. In *Theorizing the Americanist Tradition*, ed. by Lisa Philips Valentine and Regna Darnell, 207–226. Toronto: University of Toronto Press.

Rudes, Blair A. 2002. Early Vocabulary and Dictionary Development. A Cautionary Note. In Indigenous Languages Across the Community ed. by Barbara Bunaby and Jon Reyhner, 189–195. Flagstaff, AZ: Northen Arizona University. Online: http://jan.ucc.nau.edu/~jar/ILAC/ILAC_20.pdf

Ruz, Mario Humberto, and Diana Birrichaga, eds. 1997. *Las lenguas del Chiapas Colonial. Manuscritos en la biblioteca nacional de París*. Vol. dos: lengua zoque. Mexico City: Universidad Nacional Autónoma de México.

Schneider, Cindy. 2009. *Why field linguists should pay more attention to applied linguistic research*. Paper presented at the Australian Linguistic Society Conference, Melbourne, July 2009.

Siebert, Frank T., Jr. 1975. Resurrecting Virginia Algonquian from the dead: The reconstituted and historical phonology of Powhatan. In *Studies in Southeastern Indian languages*, ed. by James M. Crawford, 285–453. Athens, Greece: The University of Georgia Press.

Sims, Christine, and Hilaire Valiquette. 2000. Wick Miller's Acoma Keresan Work: An assessment. In *Uto-Aztecan: Structural, Temporal, and Geographic Perspectives. Papers in Memory of Wick R. Miller by the Friends of Uto-Aztecan*, ed. by Eugene H. Casad and Thomas L. Willett, 19–32. Hermosillo, Sonora: Editorial Unison.

Sommerfelt, Alf. 1978. *Le Breton parlé à Saint-Pol-de –Léon. Phonétique et morphologie*. New edition by F. Falc'hun and Magne Oftedal. (Monographs in Celtic Studies from the University of Oslo 1.) Oslo-Bergen-Tromsø: Universitetsforlaget.

Stearns, MacDonald, Jr. 1978. *Crimean Gothic*. Analysis and Etymology of the Corpus. Saratoga, CA: Anma Libri.

Thieberger, Nicholas, ed. 1995. *Paper and Talk: A manual for Reconstituting materials in Australian Indigenous Languages from Historical Sources*. Canberra: Aboriginal Studies Press.

Troike, Rudolph C. 1959. A Descriptive Phonology and Morphology of Coahuilteco. Unpublished Ph.D. dissertation, University of Texas at Austin, TX.

Troike, Rudolph C. 1981. Subject-object concord in Coahuilteco. *Language* 57(3):658–673.

Troike, Rudolph C. 1996. Sketch of Coahuilteco, a Language Isolate of Texas. In *Handbook of North American Indians*, Vol. 17, Languages, ed. by Ives Goddard, 644–665. Washington, D.C.: Smithsonian Institution.

Tuttle, Siri G. 2003. Archival phonetics: tone and stress in Tanana Athabaskan. *Anthropological Linguistics* 45(3):316–336.

Voegelin, Carl F. 1954. A modern method for field work treatment of previously collected texts. *Journal of American Folklore* 67:15–20.

Voorhis, Paul. 1996. Analysis of prephonemic texts: Frank Speck's Catawba, Mohegan, Penobscot, and Wawenock. In *nikotwâsik iskwâhtêm, pâskitêpayih! Studies in Honour of H.C. Wolfart*, ed. by John D. Nichols and Arden C. Ogg, 461–490. (Algonquian and Iroquoian Linguistics Memoir 13.) Winnipeg, Manitoba: Algonquian and Iroquoian Linguistics.

Warner, Natasha, Quirina Luna, and Lynnika Butler. 2007. Ethics and revitalization of dormant languages: The Mutsun language. *Language Documentation and Conservation* 1(1):58–76. Online: http://nflrc.hawaii.edu/ldc/

Werner, David. 1977. *Where There Is No Doctor; a Village Health Care Handbook.* Palo Alto, CA: Hesperian Foundation.

Wmffre, Iwan. 1998. *Late Cornish*. (Languages of the World/Materials 135.) Munich and Newcastle: Lincom Europa.

Wolfart, H. Christoph. 1999. Authenticity and *Aggiornamiento* in Spoken texts and their Critical Edition. In *Theorizing the Americanist Tradition*, ed. by Lisa Philips Valentine and Regna Darnell, 121–148. Toronto: University of Toronto Press.

Zamponi, Raoul. 2004. Fragments of Waikuri (Baja California). *Anthropological Linguistics* 46(2):156–193.

Chapter 6
Fieldwork Ethics: The Rights and Responsibilities of the Fieldworker

6.1 Introduction

There has been a great deal of discussion in recent years on the responsibilities of the fieldworker with respect to the community whose language is being studied. As Dwyer (2006:50) puts it, “The ethical requirements of fieldwork-based investigation are complex, as they demand that the researcher attend both to a respectful and reciprocal relationship with the language community and produce documentation meeting the standards of the academic community and the funding agency.” In this chapter, we review what has been reported in the literature regarding the challenges of meeting these many demands.

Topics discussed include the accurate and timely collection, description, and archiving of data; advocating for, empowering, and mobilizing community members to be effective agents of language and cultural maintenance; exercising proper caution and providing appropriate guidance for students sent to the field; appropriately attributing data sources, acknowledging/honoring data ownership, and establishing and maintaining appropriate safeguards against disallowed access to data; being aware of the consequences of fieldwork, many of which are unintended; and personal behavior in the field, including relations with consultants. Many of the ideas found in this chapter are based on our reading of Bowern (2008), Childs (2007), Crowley (2007), Grinevald (2006), Newman and Ratliff (2001), Rice (2006), and Dwyer (2006).

6.2 Documentation

As discussed in Ladefoged (1992) and Dwyer (2006:33), not everyone agrees that endangered language documentation should be a primary concern for social scientists, since communities themselves might be perfectly happy to shift to another language or might be only ‘mildly sad’ at their language becoming moribund

S.L. Chelliah and W.J. de Reuse, *Handbook of Descriptive Linguistic Fieldwork*,
DOI 10.1007/978-90-481-9026-3_6,

(Crowley 2007:51).[1] To many linguists, however, one of the primary responsibilities of a field linguist is to assist communities with language preservation and revitalization efforts through fieldwork, documentation, and description. As discussed in Hale (1992a, 1992b), Woodbury (1993), Bradley and Bradley (2002:xi–xii), and Himmelmann (2006:5–6), linguistic fieldworkers should use their training to produce materials such as dictionaries, text collections, and grammars of endangered languages so that communities can regain or maintain access to inherited knowledge.[2]

The reasons this should be done have been repeated often. When speakers lose ways to talk about their material culture, they also lose knowledge of the culture itself. For example, if a community no longer recalls the names of indigenous plants, they may also lose their knowledge of how such plants can be used for medicinal purposes. The same is true for religious rituals and other culturally specific celebrations that cannot be performed without the accompanying ritualistic language (Harrison 2007). Also, while it is unclear whether social problems such as alcoholism can be directly attributed to language loss, it is interesting to note, as discussed in Warner et al. (2007:72) and Hinton (2003), that language revival programs correlate with increased self-esteem and affirmation of identity. An excellent example of this is the Klallam Language Program where community members see a connection between increased self-esteem and an active language revival program, as described by Timothy Montler:

> (...) the success of the language program is reflected in the improved scores of Native American students on the WASL, the standard achievement test required in State of Washington public schools. In the five years since the establishment of Klallam language classes at the high school, the scores for Native American students have risen at a faster rate than in any other segment of the school population. Community members attribute this rise to the fact that the children in high school these past eight years are the first Klallam generation to have been explicitly taught their native culture and language. They are the first generation to be shown that they are heirs to a deeply elegant, complex, and beautiful heritage. The Klallam Language Program counteracts several generations of shame and humiliation and contributes to the Klallam people feeling that they deserve to conduct their lives with pride and dignity. The fact that the language is also taught to European Americans in high school has furthermore contributed to increasing mutual respect and breaking down barriers of racism (Montler 2007).

The field linguist may also feel a responsibility to increase the availability of well-analyzed, typologically unusual data, since these lead to more accurate linguistic typologies, linguistic theories and historical reconstructions. The typologist in us is fascinated with the linguistic treasures held in small languages; as Balthazar Bickel (p.c.) puts it, universals are interesting, but the excitement of doing fieldwork lies in "discovering the diversity in language."

---

[1] If communities have this view of the loss of their language, it is more than likely because generations of speakers have been culturally and linguistically stigmatized. Speakers may be willing to trade away their linguistic heritage rather than continue facing oppression. Also, as discussed in Dorian (1993), speakers may feel a strong sense of loss once the language is gone.

[2] See the UNESCO Convention for the Safeguarding of Intangible Cultural Heritage http://www.Unesco.org/culture/ich.

6.3 Responsibility for Speakers: Advocating, Empowering, Mobilizing

In earlier views, a fieldworker was thought to be fulfilling his/her responsibilities to a particular community as long as speakers were unharmed and were appropriately compensated for their time (Rice 2006:3). This way of looking at fieldwork has been superseded by consideration of the consequences of linguistic fieldwork for the community and for the individual speakers studied. It is felt that the fieldworker must use whatever s/he knows and has learned to the benefit the studied community. Labov (1982) calls this THE PRINCIPLE OF DEBT INCURRED. Cameron et al. (1992:22) further extend the linguist's responsibility, stating that the linguist's expertise should go as far as it can in supporting community needs. The fieldworker may promote language maintenance and revitalization through creating pedagogical materials, setting up language programs, discussing language planning policies with local authorities, supporting existing language maintenance programs, and encouraging new language maintenance programs. As discussed in Crowley (2007:48–49) and Ostler (1998), with programmatic suggestions in Nathan (2006:364–365) and Childs (2007), linguists should deliver usable materials. They should integrate linguistic documentation and speaker needs to deliver products that can be used to counteract language endangerment.

Finally, it is suggested in the literature that a fieldworker should empower speakers so that they can meet their own linguistic needs. The fieldworker should train interested speakers in the research-related activities that speakers themselves see as most important, utilizing methods that they deem to be most appropriate in language issues (Cameron 1998: 22–24 and Grinevald 1998:157–159). We can contrast this to Hale's approach, in which native speakers are trained to think like linguists since, in that view, the linguist's goal is to gain useful insights into grammatical structure (Simpson et al. 2001:xiv). Facilitating speaker attendance at language documentation and maintenance conferences will also benefit language activists (Crowley 2007). Specific case studies of linguists' involvement in language activism can be found in Penfield et al. (2008) and Florey (2008).

One way of ensuring community empowerment is to seek community involvement in all levels of language investigation. This 'cooperative' model of linguistic research would, for example, allow for community members to set the agenda for what data should be collected and what the primary objectives of language analysis should be.

Individual researchers must decide whether and to what extent cooperative research is possible in their particular situations. What we can see by reading the language documentation case studies is that cooperative research is reported to be personally rewarding; it enriches the goals of the linguist, and it can benefit the community in a sustainable way (Rice 2006; Yamada 2007; Czaykowska-Higgins 2009). However, a linguist needs to be prepared for success or failure in achieving cooperative goals, since so many factors are out of the researcher's control. See, for example, Childs' (2007) account of his team's documentation project of Krim and

Bom, two languages spoken in southern Sierra Leone. He explains that "however well intentioned [their plans] originally were, they often were frustrated or even thwarted in realization" by, for example, theft of supplies, and withdrawal of an offer to stay for free at the university hostel. In cooperative fieldwork, the researcher must also be prepared to give up the power to control sessions and outcomes. Additionally, the fieldworker may experience a sense of insecurity and failure because co-operative fieldwork entails being able to understand and produce the communicative routines appropriate to community norms, which will take some time to learn. (See for example Riley's (2009) experiences as an ethnographic fieldworker in France.)

Another factor in reaching cooperative goals is the fieldworker's assuming the role of "expert in the community's language". The fieldworker needs to tread softly when sharing expertise. Sadaf Munshi reports, that for her fieldwork in India, it was not appropriate for her to be the outsider know-it-all, especially since she is female and at the time was under 35 (Munshi, p.c.). Linguists should provide advice when asked. Similar reflection in Ahlers (2009:238) is based on her work on Elem Pomo, where one of the speakers expressed admiration for a linguist's knowledge of the language but felt "intimidated" by the expression of that knowledge.

An interesting question in starting a community-based project is defining what 'community' means. As discussed in Holton (2009:169), it is not easy to say who represents the community: is it a regional language center, a tribal council, a set of elders, or one particular interested speaker? It takes, in Holton's words, 'a significant amount of finesse on the part of the researcher,' to determine who speaks for the community and its needs. Working with the Lamkang of Manipur state, India, I [Chelliah], am aware that each village has a chief, that those chiefs change by a regular election process, and that each chief has his own concerns for the language. Thus working cooperatively with one chief does not necessarily ensure meeting the concerns of other chiefs and their respective villages. Furthermore, village leadership may not represent a community's concerns. As described by Holton (2009), in his work in Eastern Indonesia, many co-operative projects begin from the bottom-up. In his case, Holton's collaborator was interested in developing a dictionary. As fieldworker and consultant began to collect words, village experts in different technical areas were consulted for input, and the number of speakers involved in the project increased so that 'everyone in the community was soon aware that a dictionary of Western Pantar was in the making…' and everyone felt free to contribute (171). Thus this project became a community-based project even though it did not start as one (172).

Linguistic fieldwork leads to the creation of different products, and these products may fulfill one or many goals. One goal of linguistic fieldwork, for instance, is to create a record of the language for use by linguists and speakers. A typical product is a linguistic grammar. Another goal may be language revitalization, for which relevant products may be an annotated text collection or pedagogical materials. In fact, language teaching and learning on the one hand, and language documentation and description on the other, are synergistic activities. A report of these complementary activities can be found in Francis and Gómez (2009) for work on Nahuatl spoken by the indigenous communities of the Malintzin volcano highlands in Mexico.

Often times a funding agency determines what the goals will be. Dwyer (2006:36) notes that, "If a researcher is funded by a university in that nation's capital, for example, in some cases he/she might be expected to produce a study that enhanced that country's ethnic policy." With regard to local governments, the ethics page of the American Anthropological Association website,[3] recommends that a fieldworker should "demand assurance that they will not be required to compromise their professional responsibilities and ethics as a condition of their permission to pursue research. Specifically, no secret research, no secret reports or debriefings of any kind should be agreed to or given. If these matters are clearly understood in advance, serious complications and misunderstandings can generally be avoided."

It is natural that researchers tend to privilege those goals that they can easily implement and that have the best consequences for their academic careers. Linguists have the resources in terms of training and equipment, as well as funding and institutional support, to work for extended periods of time on linguistic descriptive projects that result in publications. But only some publications will fit in with the goals of the community. So, how does a linguist look past his or her goals and discover the goals of a community? The best approach is to ask. There are usually some community members who stand out as interested in the study of language or folklore, and they might be a window to the needs of the community. Some common goals are language revitalization, revival, or documentation. To academics, the route to revitalization and maintenance is tied to literacy. It should not, however, be blindly assumed that every community wants their language to be written down. Some speakers will resist having their language written down, motivated perhaps by fears that, once written down, texts which should have restricted access could be circulated and shared outside of the community. In this case, the linguist needs to tease out community goals to see how one set of goals (language maintenance through increased literacy, for example) can be pursued without hurting other goals (preserving the integrity of sacred texts, for example). Debenport (2010) reports on such activity in a Pueblo community of New Mexico, where speakers met to review written materials proposed for inclusion in the community's dictionary; each word and illustrative sentence was vetted for culturally sensitive material. Thus the dictionary was created while honoring "local language ideologies regarding textual circulation and materiality" (Debenport 2010).

One question that researchers can ask themselves is whether they are trying to do too much. For example, should a linguist take on the very important tasks of creating pedagogical materials, designing curriculum, training language teachers, and/or designing and producing classroom materials? Most linguists have no training in language teaching, and should seek the help of trained specialists or take courses on these methods.[4] It would be unethical to waste precious time

[3] http://www.aaanet.org/stmts/ethstmnt.htm

[4] These concerns are being addressed by the field as evidenced by the recent symposium Fostering Synergistic Partnerships between Teachers and Linguists, held at the Linguistics Society of America Annual meeting, San Francisco, January 8, 2009. See also Schneider (2009) and von Gleich (2005) for suggestions on developing language teaching materials in documentation projects.

producing second-rate materials that are ineffective in revitalization or stabilization efforts. Project teams with diverse expertise can successfully fulfill multiple goals (see Section 8.7) without compromising product quality and outcomes.

Not all languages on which linguists conduct descriptive fieldwork are endangered, so these speech communities will have different goals. In my [Chelliah's] work on Manipuri, which has more than a million speakers, interested community members often have degrees in Linguistics, and may themselves be language educators or activists. Their goals are to work towards spelling standardization through the creation of dictionaries; to raise the international status of the language through Unicode font development; to resolve complex orthographic choices; to secure steady jobs for those trained in linguistics; to develop a protocol for morphological tagging; and to secure more advanced linguistic training and help in publishing in international venues. Empowering community members to reach these goals takes long-term community involvement. Notice that the products of the goals expressed by the Manipuri linguists are not easily assessed for rewards in the American academic context. A book, an article, or funded grants are easily awarded brownie points towards tenure or promotion and raises. But it is not obvious how research activities such as "trained linguists"; "proofread and revised article" or "helped with Unicode issues" will be evaluated. It is a challenge that our current academic system of research evaluation needs to meet.

Keeping the goals of the community paramount is not always easy. At times a linguist's enthusiasm for a linguistic revelation may be at odds with the concerns of the speakers are at odds. For example, it has been our experience that speakers of some north-east Indian Tibeto-Burman languages object to their language being classified under a particular subgroup. While we, as linguists, might find solid and even exciting confirmation of a language's genetic placement, we find that we must respect the culturally and politically based resistance to our suggested classification. The fact of the matter is that academic concerns can be strong. Sara Hale reports that Ken Hale was "…working on Damin, the ceremonial language, which he was very excited about. He was sworn to secrecy about it, and he kept the secret pretty well for quite a while. But then it just became too fantastic because he was very excited about the insight that you got from Damin into Aboriginal semantics…" (Hale 2001:26).

There can also be a dissonance between academic and community concerns and the rhetoric that is used to express these concerns. For instance, as discussed in Hill (2002), books and articles on documentary linguistics characterize all languages as priceless resources that enrich the knowledge systems of the world and therefore need to be recorded and analyzed. Characterizing a language as a priceless resource may contradict speakers' aims of regaining the use of their language as a tool for day-to-day communication. Similarly, England (2002:141) points out that we characterize languages as endangered without considering that this characterization might make speakers feel even more marginalized. It represents their language as "suffering loss" rather than "stable and increasingly expanding." Thus, as Hill suggests, advocacy projects should be carefully crafted so that the rhetoric highlights speakers' attitudes towards their language as well as the documentation project.

6.4 Responsibility for Students

Advisors are responsible for students who undertake fieldwork as research assistants or Ph.D. candidates. The Linguistic Society of America's Ethics committee[5] states that faculty should: ensure the safety of students; exhibit ethical behavior and thus encourage the same in their students; duly acknowledge contributions of the student researcher; support and supervise student research; avoid sexual liaisons with students; and instruct students in what constitutes ethical behavior for the field. While these guidelines seemingly state the obvious, it is not clear how often faculty routinely and openly discuss these issues with their student fieldworkers. Graduate student mentoring is not part of academic professional training and most academics are not prepared for the time and effort it takes to appropriately mentor students.

The Research Center for Linguistic Typology at La Trobe University has put out a fieldwork manual, *Fieldwork and Your Wellbeing* (2009), which discusses issues of field safety and field preparation in the field some detail. It would be useful to require students to read a manual of this sort, or to read Chapter 5 and 6 of this book. Faculty and students should deliberate on these issues as a matter of course.

6.5 Honesty in Research

At the 41th meeting of the International Conference on Sino-Tibetan Linguistics and Languages held in London, a researcher declared, "It dawned on me a week ago at dawn that I've been wrong all these years [about a point of grammar]." The motivation behind this honest declaration is one that should be emulated. Incorrect analyses can slow down progress in description. For example, Krauss (2005:61) characterizes the effect of Edward Sapir's unwillingness to admit mistakes in the analysis of Athabascan tone as setting back "…the development of comparative Athabaskan syllable nuclei by four decades." Furthermore, since the linguist "in charge of" the description of a language is looked to as the authority on that language by the academic community and often by the speakers as well, and since there are real linguistic consequences of description – analyses filtering into pedagogy, and therefore becoming a part of the record used for revitalization efforts – it is the responsibility of the researcher to revise results as necessary rather than sweep them under the rug.

Needless to say, it is unethical to fabricate results. But it is equally unethical to assert an analysis based on data from a single speaker when that data has not been

[5] http://lsaethics.wordpress.com/category/ethics-statement/

replicated with other speakers or supported with data from texts. One way of ethically reporting results of data elicited from a single speaker is to clearly state that the data source is limited to one speaker, thereby allowing the reader to properly assess the strength of the data for the analysis being put forward. There is, no doubt, value in single-speaker data, as this may be the only data source available in the case of an endangered language. However, we support the view that speaker data backed up by replication in texts and the speech of other speakers strengthens the validity of a data set, illustrating that elicitation effects have not shaped the resulting data set.

Another point regarding academic ethics has to do with the pace of producing usable material. It takes little time to record a narrative, but it might take months or even years to do a good job of transcribing or translating that narrative. Similarly, it takes years to complete a grammar or dictionary of a language. This we know. On the other hand are the valid complaints from Native American language activists that, in the past, linguists have 'sat on' data for years while a language quickly faded away. It has become increasingly clear that, rather than interminably incubate the data egg, anticipating the hatching of an impressive grammar or dictionary, the linguist must produce, with the assistance of local and non-local experts, smaller intermediary user friendly products such as grammatical sketches, illustrated dictionaries based on semantic fields, pedagogical materials and the like. (See Section 6.3 for a list of such products.) In the case of an endangered language, it would be unethical to record data but not process it further (Bowern 2008). Paul Newman reflects about linguists who conduct empirical linguistic research:

> [we]…have decried the dominance of linguistic theory in our discipline, but we have tended to remain silent about an equally serious weakness affecting the descriptive linguistic enterprise, namely the phenomenon of field linguists who fail to write up and publish their findings. All of us, whether specialists in Africa or Southeast Asia or Latin America, know of legendary figures –whom we usually mention in reverential terms – who have mountains of knowledge in their heads and masses of materials in their files but who have published very little. These materials cry out for both readings of the C in LD&C [Language Documentation and Conservation], conservation and communication, with the latter being as urgent as the former. (Newman 2007:28)

It seems reasonable to say that if fieldworkers have access to speakers, they should complete as much analysis as possible even as they continue to gather data so that their work is of use to others. However, since academics live under the rule of publish or perish, there is a constant need to publish. Community-accessible products, such as semantic-field dictionaries and pedagogical materials, are usually not peer-reviewed before publication and therefore do not "count" towards career advancement. Also not peer-reviewed are online archives. A well-done archive – with analyzed texts, sound files, grammatical descriptions, and cultural information – is a treasure trove for linguists of every ilk and also for community members. It takes years to build, and requires the effort of several participants. As of yet, however, there are no mechanisms in place to peer-review online archives, so these, too, are problematic for academic institutions to take into account for tenure or promotion.

Here, then, is the unsolved contradiction. On the one hand, academic institutions crave all the benefits of administering a funded project. On the other hand, there are no mechanisms in place to give credit for the products that those grants require. The Linguistic Society of America has proposed that non-traditional publications should be valued as highly as traditional publications, so that linguists can fulfill their ethical obligations to themselves, their funding agencies, and the communities they work with.[6] It is yet to be seen how promotion and tenure committees and merit committees implement the assessment of these non-traditional scholarly works and determine what value they have for the granting of tenure, promotion, or merit raises.

Crowley (2007:52–56) discusses some common-sense rules of etiquette for linguistic fieldworkers that should be considered before even selecting a language to work on. It does not pay to be territorial about the language one works on. If community members welcome another linguist to work on their language to supplement existing research, so be it. There is plenty of work for everyone. On the other hand, if a linguist is looking for a language to work on, he or she should try not to step on anyone else's toes or to ruin a good arrangement that another linguist has established with a community. If it is absolutely necessary to share a field site, researchers should be: (1) careful not to bad-mouth each other and (2) choose a slightly different focus of investigation from ongoing research in the community; and (3) share data, consultants, resources, and products whenever possible.

6.6 Data Attribution and Ownership

Related to the idea of honesty in research is proper attribution of data and analyses. There are clear guidelines for this as far as published academic work is concerned (e.g. Raign 2006:463–479). For less universally understood copyright laws for academic publications – such as cases of joint or group authorship – see Newman (2007).

Regarding data attribution in publications and other products like websites that are the result of linguistic fieldwork, it has become commonplace to find each data example tagged with the name of the speaker, as in, for example, Mithun (2001). Minimally, all participants and contributors – with details of the extent of their involvement in the project – should be acknowledged. Of course, there are instances where speakers do not want their names mentioned. Permissions to use speaker names should be taken at the time of recording. What if someone other than the

[6] We are referring to The Linguistic Society of America resolution recognizing the scholarly merit of language documentation which was prepared by the LSA's Committee on Endangered Languages and their Preservation. The resolution can be found online at http://www.lsadc.org/info/lsa-res-lang-doc.cfm).

narrator is mentioned by name in a recording? Dwyer (2006:48) states that this person, too, should be "involved in decisions of access." If there are research helpers, they too should be consulted before being named in any publication. Finally, Dwyer advises not attributing utterances to particular persons when, "villagers allow full access including crediting recordings to their name, but local coordinators, possessing an overview of social issues, suggest anonymity for political reasons."

Next, the fieldworker should consider the question of data ownership. Copyright rules determine who legally owns a particular product, and therefore who can sell or make a profit from that product. Copyright rules differ from country to country; what we are discussing here is specific to U.S. copyright law. What of copyright of oral material recorded in the field, such as a narrative? Newman (2007:40) states that, for oral material, no one legally owns the copyright until it is recorded on paper or other media. Once recorded, the copyright is "automatically attached", and whatever the researcher intends to do with the material must be approved by the speaker. It is best to get these approvals at the time of recording and, if at all possible, to remind speakers of their involvement in the project before the release of those materials. For example, if materials are to be published on a website with annotated texts and associated sound files, speakers could be given access to preview the website before public release.

It is important to note that payment does not necessarily imply ownership. Newman reports that when the fieldworker works with a speaker to translate a text, just because s/he has paid the translator for his or her time, does not necessarily mean that the fieldworker owns the resulting translation (Newman 2007:41). There must have been an agreement of "work for hire" for that to be true. If such an agreement has not been signed, then the researcher owns 50% interest, and the translator owns the remaining 50% of the translation. Any royalties that arise from resulting products would be the property of both the translator and researcher. This is obviously a very complicated situation, as there can be several people who work to bring a text to a publishable stage.

It makes sense, in terms of maintaining positive and respectful speaker-fieldworker relationships, to acknowledge the moral rights of the speaker (Dwyer 2006:48). The Ethics page of Max Planck Institute of Evolutionary Anthropology[7] instructs anthropologists and linguists to follow these rules:

> Especially given the increasing importance of intellectual and cultural property rights, individuals or communities participating in research should be informed that the institute and the researcher seek the right to store, use, and disseminate (with restrictions where appropriate) the material in question, but do not assert ownership of the intellectual or cultural materials entrusted to the institute or the researcher. When stored and disseminated, such materials should always make due acknowledgement to their authors and performers. Authors / performers should be named explicitly only where their informed consent to this has been obtained; otherwise, an anonymous acknowledgement is appropriate. It is appropriate for the researcher to pay the individuals

[7] http://www.eva.mpg.de/lingua/resources/ethics.php

involved in research for their time and travel and other out-of-pocket expenses. It is not appropriate to make payments that might be construed as payments for the transfer of ownership.

6.7 Rights and Responsibilities Regarding Access to Data

With access to data comes the need to share those data responsibly. First, fieldworkers must think of how to make their findings available to other researchers, the idea being that many eyes on the same data can lead to more discoveries. We have found that fieldworkers are not unequivocally happy about making their data and field notes available with no restrictions to the academic community. It is a little like buying a jigsaw puzzle, and then being asked to hand it over to someone else who gets the enjoyment and credit for figuring out how all the pieces fit together. Since, as we have been maintaining all along in this book, descriptive linguistic fieldwork is both data collection and analysis, a field worker might rightfully assert that since they have worked on the data from the outset, s/he knows the data best and is thus best able to complete at least a first round of analysis.

While the data do not belong to an individual researcher (see Section 6.6 on copyright), s/he has control over the products of fieldwork, for example, field recordings; field notes; time-aligned annotations of texts; word-for-word analyses; full interlinear glossing of texts, with morphological analysis; free translations; word lists; cultural notes; analytic notes; pictures; songs; or scans of printed materials. The fieldworker should responsibly decide how much, when, and who should have access to the data s/he has spent time collecting. These decisions should be made keeping in mind the needs of the speakers, the funding agency, the success of the fieldworker's project, and other researchers in the field. What an academician thinks of as "protecting" data could be considered "hoarding" by community members who have urgent need of some of the products. Take, for example, the words of Andrew Balluta, a speaker of Dena'ina (Alaska), expressing his community's needs for access to recordings:

> You know, all these recordings … if we don't get it out and learn about it, where are we going to learn from? These are old recordings. We want to get it out and teach our younger children what the elder people are talking about. I think that's a very good idea for getting it free so we can listen to them. (Holton 2005)

As discussed in Section 7.8, field materials should be archived so that they are in a safe and accessible place for future researchers. A powerful argument for preservation comes from Ellen Demit, a speaker of Tanacross (Healy Lake, Alaska) who says to the linguists working on her language, "You guys I want to put this in your heart, in your mind. That one day—how many recordings I make for you? Respect, take care. Don't say "I lost that recording." My words are very strong words, you gotta keep that recording." (Holton 2005). A digital archive is preferable to hard copy because it allows the researcher to have all the materials in one place with different levels of access set for individual data. A good example of a graded access system is The Archive of Indigenous languages of Latin America (AILLA) based at the University of Texas at Austin (Dwyer and Mosel 2001; Dwyer 2006). This archive sets data access restrictions so that some parts of the archive may be fully closed, some open only to particular archive users such as team researchers or specific speakers or community members, and some open to all.

With the ability to restrict access, a researcher can allow certain texts to be viewed and/or downloaded and used, but other texts to be restricted completely. This would be useful in a scenario discussed by Bowern (2008:188), where there are some previously published texts that the researcher has used for analysis but does not want to disseminate through the archive because some speakers object to the texts being published in the first place. Sacred, secret, or sensitive information should obviously have restricted access.

In addition to access restrictions based on speaker expectations, requests, and permissions, the researcher can also consider how to use graded access to release field notes. For example, it would be reasonable to provide access to as many time-aligned annotations or interlinear analyses as possible because these may be urgently needed by community members for creating pedagogical materials. Field notes and analyses can be made available with rules about how these can be used and cited, such as: 'permission to quote but not to copy'; 'copying permitted for private study only'; 'no quotation allowed without permission of the author' (Bowern 2008:186). In this way, a researcher can make original data available with the same expectations of being given credit as with paper publications. It is important for contributors to closely monitor access restrictions. Seeger (2001) reports that many collections in the Indiana University Archives of Traditional Music had been deposited with complete restrictions on the whole deposit because, "the researchers who made them wanted to publish their results before others could use their collections. They would restrict the collection, and then forget to change it after they had published their results. As part of a broad effort to improve access, [Seeger] contacted every depositor [he] could find to renegotiate their contracts." (Seeger 2001)

Ultimately, it is speakers who decide what material can be disseminated. However, not all speakers are equally responsible or sensitive to their own community's wishes. For example, Innes (2010) reports that when doing philological work on a body of texts collected by Mary Haas on Muskogee (Creek), she found that speakers felt uncomfortable with many of the texts because of their power. Contact with the texts was deemed emotionally and physically dangerous. By putting some of these stories into writing, it was felt that the power of the texts was strengthened. It is not clear why or under what circumstances the narrator shared these texts with Mary Haas. What is clear is that subsequent sensibilities did not allow for the texts to be openly shared. As discussed in Section 6.8, fieldworkers cannot always be sure that speakers themselves realize the implications of the data they are sharing, even though they have given their informed consent. They may, for example, have no understanding of how widely their information can be accessed through web dissemination, or even what the Internet is. In these instances, the fieldworker has a moral obligation to respect not only the letter of agreement (the signed consent form), but the spirit of the agreement. For instance, Robinson (2010) notes that her Dupaningan Agta (Philippine) consultants do not know what web dissemination implies; but, since they have given permission for a particular language resource to be viewed 'by everyone' she feels that wide distribution is alright.

If a fieldworker does not intend to work on a language any longer, it is important for him or her to pass the torch to another linguist who is interested in continuing the work. It is a little tricky to see how permissions given to one fieldworker can be transmitted to another fieldworker as communities may have given permissions to a particular fieldworker with whom they had developed a relationship. To protect sensitive information, the fieldworker should only designate control of the archive to a researcher who has high standards in preserving access restrictions and who has community approval.

Finally, with access to data and with the training and motivation to analyze the data, come responsibilities to provide speakers with materials which are useful for language revitalization and stabilization. Some common products that can be used by speakers are a script and spelling system, pedagogical grammars, phrasebooks,[8] learners' dictionaries, primers and other reading materials, and other electronic and print pedagogical materials. As these are not directly related to the descriptive linguistic enterprise, we will not discuss them in detail; the reader may refer to relevant discussions in the ever-growing literature on endangered language documentation, some of which is listed at the end of this chapter.

The materials produced by speakers should be intellectually accessible to speakers. In the typical instance, a reference grammar will form the basis for more user-friendly pedagogical manuals of the language. Reference grammars are usually written for other linguists; they are analytic, grammatical description is based on form rather than function, and require some prior knowledge of terminology (Evans and Dench 2006:14–15). On the basis of reference grammars, different grammatical treatments can be produced – a comprehensive descriptive grammar, a textbook designed for classroom use, or a learner's manual – based on communicative needs. For English, for example, we have the well-known grammars by Quirk et al. (1985), which is encyclopedic, and Azar (1989), which is a classroom textbook with exercises. These are organized by units such as tense, aspect, imperative and question formation. The Collins Cobuild grammar (Bullon et al. 1990), on the other hand, is functionally oriented, and organized by units such as "expressing time".

Whatever else is done, it is important to make materials available to speakers on an ongoing basis. Crowley (2007:34) notes that some countries have closed off areas to foreign researchers because previous fieldworkers had not shared research results.

---

[8] Some examples of pedagogical grammars are mentioned in Section 11.2.1.2, and some examples of conversational phrasebooks are in Section 12.2.2.6.

6.8 Unintended Consequences of Fieldwork

There are certain adverse consequences of fieldwork that fieldworkers should try to minimize. One obvious effect is that by working with specific individuals in a community, the fieldworker changes their lives by changing their economic or power status in the community. The fieldworker should be careful to not make consultants the target of community jealousy. This can be done by giving others ample opportunities to participate in the fieldwork project and by not advertising payment amounts unnecessarily.

There are some unforeseen effects that should be guarded against. In many cases, speakers of endangered languages are in dialogue with governments on issues such as land or water rights, linguistic or political freedom, religious support of one or another mode of behavior, or the selection of one script over another. If the fieldworker collects narratives that support one interpretation of the community's history over another, there could be serious consequences for ongoing dialogue with government agencies. Therefore, this kind of text should not be recorded or at least should not be disseminated. Crowley says of narratives, "You should be aware…that what may appear to be an innocent children's tale…could be construed … [as] a claim for land based on the speaker's knowledge of the story and that story's association with spirits inhabiting certain areas of land." (Crowley 2007:25). He continues with the following example: "…a story about a family that is descended from some particular animal might be little more than a story, [but] it may well be that this story is implicitly taken by local people to indicate that the family involved has rights to land in whatever places are associated with the … [animal] in the story." (Crowley 2007:26). Eira (2009:309) also emphasizes this point in her review of Crowley (2007).

Another point to consider is the potential danger of disseminating analyzed stories or conversations which belittle a member of the community. If the community member is powerful, this might cause serious problems for the consultant. Personal narratives might also exaggerate the heroics of one character and the bad behavior of another. At times, these stories might be out-and-out-lies. Such stories would in effect be libelous towards individual community members, and by extension towards their families. So, even if the narratives make for good linguistic data, they cannot be used as examples in a corpus (Crowley 2007:26).

One of the desired outcomes of language documentation and description is a renewed and revitalized interest of speakers in their language. There are certain consequences to this revitalization that fieldworkers should keep in mind. For example, innovated features may be inadvertently introduced into a revived variety when revitalization is attempted with previously collected language data, but with no native speaker input (Warner et al. 2007). Is it ethical to say one is revitalizing one variety but actually create a different variety? Also, consider the case of smaller speech communities in Northeast India whose unique linguistic identity is precariously balanced against the linguistic identity of several other small languages. In this case, documenting one language while ignoring the others will set the existing ecology off balance, perhaps to the detriment of the very community the linguist is trying to help.[9] Similarly, humanitarian aid projects supported by linguists, as discussed in Epps and Ladley (2009: 645), can also upset the balance of local power by, for instance, increasing resources by providing well water in one area but not another.

In India, if a language has a script and pedagogical material, it can lobby for government support in the form of funds for language development. Linguists should keep in mind that language can be used as political symbol to motivate what Annamalai (1998:26) calls a “reversal of assimilation”. He cites the case of the Bodo of Assam (Northeast India) who were assimilating with the dominant Assamese population, but for whom “the path of assimilation started to be reversed with concerted efforts for teaching Bodo in school”. Violence and political instability in Assam have followed, as the Bodo attempt to gain political and administrative autonomy distinguishing them from the majority population. Of course, there is no reason to believe that language maintenance or revival will be the sole cause or a predictable cause of civil unrest, but the fieldworker needs to be aware that any materials which provide a unique identity to a community can be used as political symbols. That is why language planning should always be conducted in conjunction with experts in language planning and policy.

6.9 Christian Evangelism and Fieldwork

In a recent issue of the journal *Language* (Volume 85.3, 2009), six articles discuss the motivations and outcomes of linguistic fieldwork, description, and documentation conducted by the evangelical Christian organization, the SIL International (formerly known as the Summer Institute of Linguistics). Dobrin and Good (2009: 621–628) catalog the contributions of SIL to linguistics and the communities it has served. SIL has:

- Revitalized several endangered languages in Amazonia and Melanesia
- Provided the first language descriptions for many parts of the world
- Discovered typologically unknown structures and sounds
- Developed methodologies leading to reliable data collection
- Trained teachers and created materials so that government literacy programs can succeed
- Developed and shared software for text-analysis and lexicon building

[9] Interestingly, investing time in describing typologically similar adjacent languages is regarded as low priority for typologists. When discussing how to select a language to study, Dixon advocates looking for one which is typologically interesting: an isolate, for instance, rather than a language representative of a large group where many of the languages may share typological features. (Dixon 2010:312–313). Describing a new Oceanic language when good descriptions already exist for some of the 500 languages in this group would increase our knowledge of this language family only by filling us in on a few language-specific features. In Dixon’s view, time would be better spent describing a Papuan language of the Solomon Islands, of which less is known.

- Developed and shared fonts and keyboarding tools
- Given linguists logistical assistance in the field (making contacts, selecting a field site, arranging house and transportation, teaching about the culture, and so on)
- Provided a useful institutionally-supported collaborative model for bringing about social change through activities such as language revitalization[10]
- Provided medical interventions
- Have "multifaceted relationships" with speakers that allow for the success of language programs.

In addition, Olson (2009) points to SIL's extensive library of documentation on languages. While these accomplishments are noteworthy and, in fact overlap with, the goals discussed here for descriptive linguistic fieldworkers, we are asked to consider that the ultimate goal of translating the Bible with the view to gaining coverts to Protestant Christianity is at odds with the goals of preserving cultural and linguistic diversity. Conversion to Christianity implies repudiation of traditional ways of life including "attitudes toward work and leisure, male-female relationships, use of alcohol and tobacco (whether for ritual and recreation), personal modesty, economic transactions, and so forth" (Epps and Ladley 2009:644). Christianity is seen as bringing in cultural change from the outside that results in, "the loss of linguistic styles, registers, genres, and varieties" (645) especially where those are in contest with Christian identity. Olson (2009:651) strongly contests that SIL International replaces indigenous cultures with a Western Christian culture. He argues that SIL "strengthens local cultures" and provides Christianity as a possible option. The accuracy of these two viewpoints can only be judged by taking individual language situations into account. Those faced with communities that were converted in the heyday of colonial expansion in Africa or assimilation movements in the United States could easily support the first view. Those working with communities which were converted more recently or which have robust intact populations, as in Asia and Southeast Asia, may more easily accept the second view.

A linguistic rather than cultural argument against missionary-based fieldwork is expressed by Dixon (2010), who says that missionaries who are compelled to work on Bible translation may not enjoy linguistics and may therefore avoid writing a grammar. The descriptions that result from such fieldwork will inevitably be of "low value" (2010:310).

Secular endangered language research is not a politically or culturally neutral activity, either. See Czaykowska-Higgins (2009:35–39) for an excellent exemplification of this point. Secular fieldwork is said to be based on "the new moral stance of endangered language linguistics" by which language is seen as "part of the intangible heritage of minority communities and promoted under [the belief that] humanity would be poorer for the lack of diversity that would be represented by the loss of the language" (Handman 2009:636). In this view, language, culture (and

[10] Mission projects are collective; missionary linguists work in an atmosphere of "institutional and mutual support and commitment to common goals" (Dobrin and Good 2009:627).

religion as a part of culture), are seen as a single package. According to this view, a speaker who has moved away from traditional practices would be seen as a lost cause. In our view, this would be misguided. The world has grown smaller through globalization, and the world view of literate speakers of smaller languages has grown larger. It is natural to see some decline in traditional practices because of education and advances in modern medicine. But the separation between linguistic skill and traditional practice is illustrated by the case of Western Europe, which has many fluent speakers of Romance and Germanic languages who are now atheists, having given up generations of Christian tradition!

Whether our motivations are academic, humanistic, or evangelical, we must be aware of how our language ideologies color our fieldwork activities. The ethical way to proceed is to minimize the consequences of our actions where they contradict local concerns and plans. We must trust that indigenous communities can decide what is best for themselves.

6.10 Personal Behavior

There are certain aspects of personal behavior that a fieldworker should be cautious about:

Sex: Fieldworkers have a financial advantage over their consultants. This can result in a power difference that could lead to unstated "quid pro quo" for sexual favors. Therefore, fieldworkers should be extremely careful about romantic or sexual entanglements with consultants. An anonymous blogger posted the following on the Linguistic Society of America ethics website[11]:

> There is otherwise a worrying trend that has been noticed towards a disturbingly cavalier attitude towards refraining from sexual harassment and demand for "quid pro quo" sexual favors, particularly among linguists operating in regions of the world where the usual professional and/or legal injunctions against sexual harassment and exploitation commonly observed and enforced in the United States and much of the Western world cannot be ordinarily enforced, at least without the risk of long and painfully drawn out legal process at an international level.

Religion and spirituality: Regardless of the fieldworker's belief or lack thereof in religious or spiritual matters, it is never a good idea to argue or impose one's own beliefs on a consultant. It is counterproductive to argue against the existence of ghosts if a speaker is recalling how a local haunting was remedied. When invited, it is a good idea to participate in religious ceremonies as a participant observer. Of course, there are limits to what each individual can, and perhaps even should, tolerate; for example, a fieldworker invited to attend a circumcision ceremony could politely find an excuse not to attend, if this is something that would make him/her uncomfortable.

[11] http://lsaethics.wordpress.com/2008/07/

Dress and personal appearance: There are local norms for dress and behavior. Dress is more or less an easy matter to conform to. In general, conservative clothing is universally more easily accepted than less conservative attire. Local interpretation of appropriate modesty should be followed. It is not necessarily the case that the community will approve of the fieldworker's dressing 'like a native', especially if s/he is unable to wear something like a *dhoti, sari, sarong,* or *phenek*[12] in the proper way. In India, there are several dress options, from pan-Indian to more local attire. My [Chelliah] experience has been that speakers approve of "outsiders" wearing pan-Indian clothing (a *sari*, for example) rather than attempting to assume a local identity by wearing local attire (a *phenek*, for example). In some areas, a woman's loose, long hair is frowned upon, and it should be either covered up (Islamic countries), or put up, braided, or put in a ponytail (rural Mexico). For men, beards may be frowned upon in some areas, whereas in others they confer upon the fieldworker an air of sagacity.

Smoking and drinking: In many cultures, it is perfectly acceptable for men to smoke tobacco and drink alcohol. In many parts of rural India, older women smoke or chew tobacco but younger women of "good" character are expected not to do so. The same differences for male and female behavior hold for alcohol consumption. There may be legal restrictions against alcohol; most Native American Indian reservations are dry, for example. Also, alcoholism is unfortunately common in some indigenous communities. The fieldworker should think of the consequences and implications of drinking with consultants. It is possible that community members who are trying to stay sober will not respect a fieldworker who encourages another community member to drink. Out of deference to their hard work at staying sober, the fieldworker should not drink. Conversely, in other cases, the fieldworker might be expected to drink with the community at festivals or rituals.

Straining local resources: As a guest of a particular family or community, it should be kept in mind that resources that might otherwise be used for community members are being consumed by the fieldworker. Resources include food, space (such as a bed) and goods (such as linen or mats) (Crowley 2007:173). Water and electricity should be used sparingly. The fieldworker should tactfully supplement existing resources, by, for instance, regularly buying groceries for a host family.

Usurping native standing: The fieldworker should be careful not to act like a know-it-all. By constantly knowing the answer, the linguist can unintentionally position themselves as the local language expert, thereby usurping that position from a community member. It is better for the project as a whole to have the support of the local expertise, and a local expert will be necessary for effective team fieldwork. Native speakers, especially the educated or respected members of the community, should not be corrected (Crowley 2007:175).

[12] The *phenek* is the traditional wraparound skirt worn by Manipuri women.

6.11 Conclusion

Ultimately, the most instructive way to learn about ethical behavior is to learn about the actual experiences of other fieldworkers. While anthropologists have a case book discussing real-world ethical dilemmas encountered by fieldworkers (McNamara 2010), linguists still lack a resource of this sort. For now, a useful forum for sharing field experiences on this topic is the Ethics blog on the Linguistic Society of America website at http://lsaethics.wordpress.com/.

References

Ahlers, Jocelyn C. 2009. The many meanings of collaboration: Fieldwork with the Elem Pomo. *Language and Communication* 29(3):230–243.

Annamalai, E. 1998. Language Survival in India: Challenges and Responses. In *Studies in Endangered Languages*, ed. by Kazuto Matsumura, 17–31. International Clearing House for Endangered languages Linguistic Studies 1. Tokyo: Hituzi Syobo.

Azar, Betty Schrampfer. 1989. *Understanding and Using English Grammar*. [second edition]. Upper Saddle River, NJ: Prentice Hall Regents.

Bowern, Claire. 2008. *Linguistic Fieldwork. A Practical Guide*. New York: Palgrave MacMillan.

Bradley, David, and Maya Bradley, eds. 2002. *Language Endangerment and Language Maintenance*. London: RoutledgeCurzon.

Bullon, Steve, Ramesh Khrishmurthy, Elizabeth Manning, and John Todd. 1990. *Collins Cobuild English Grammar*. New York: HarperCollins.

Cameron, Deborah, Elizabeth Frazer, Penelope Harvey, M. Ben H. Rampton, and Kay Richardson, eds. 1992. *Researching Language: Issues of Power and Method. Politics of Language*. London and New York: Routledge.

Cameron, Deborah. 1998. Problems of empowerment in linguistic research. In *Le travail du chercheur sur le terrain*. Questionner les pratiques, méthodes, les techniques de l'enquête, ed. by Mortéza Mahmoudian and Lorenza Mondada, eds., 23–38. Cahiers de l'ILSL 10. Université de Lausanne.

Childs, G. Tucker. 2007. The ethics of documenting dying languages: Lessons from the MDP and projections for the DKB. Gainesville, FL: 38th Annual Conference of African Linguistics (ACAL), March 22–25. Online: http://www.unimol.it/summerschool/teachers%20materials/TuckerChilds/Childs%202007%20Ethics.pdf

Crowley, Terry. 2007. *Field Linguistics. A Beginner's Guide*. Oxford: Oxford University Press.

Czaykowska-Higgins, Ewa. 2009. Research models, community engagement, and linguistic fieldwork: reflections on working within Canadian Indigenous Communities. *Language Documentation and Conservation* 3(1):15–50. Online: http://nflrc.hawaii.edu/ldc/http:// hdl.handle.net/10125/4423.

Debenport, Erin. 2010. The potential complexity of "universal ownership": Cultural property, textual circulation, and linguistic fieldwork. *Language and Communication* 29(4).

Dixon, R. M. W. 2010. *Basic Linguistic Theory. Volume 1 Methodology*. Oxford: Oxford University Press.

Dobrin, Lise M., and Jeff Good. 2009. Practical language development: Whose mission? *Language* 85(3):618–629.

Dorian, Nancy C. 1993. A response to Ladefoged's other view of endangered languages. *Language* 69:575–9.

Dwyer, Arienne M. 2006. Ethics and Practicalities of cooperative fieldwork and analysis. In *Essentials of Language Documentation*, ed. by Jost Gippert, Nikolaus Himmelmann and Ulrike Mosel, 31–66. Berlin and New York: Mouton de Gruyter.

Dwyer, Arienne and Ulrike Mosel. 2001. *Metadata Description Recommendations: General; Content.* DOBES Technical Reports 6a.1, 6b.2 (15.03.01).

Eira, Christina. 2009. Book reviews of: Field Linguistics, a Beginner's Guide, by Terry Crowley, and Linguistic Fieldwork: A Practical Guide, by Claire Bowern. *Australian Journal of Linguistics* 29(2):306–310.

England, Nora C. 2002. Commentary: Further rhetorical concerns. *Journal of Linguistic Anthropology* 12(2):141–143.

Epps, Patience, and Herb Ladley. 2009. Syntax, souls, or speakers? On SIL and community language development. *Language* 85(3):640–646.

Evans, Nicholas, and Alan Dench. 2006. Introduction: Catching language. In *Catching language: The Standing Challenge of Grammar Writing*. (Trends in Linguistics. Studies and Monographs 167.), ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 1–39. Berlin: Mouton de Gruyter.

Florey, Margaret. 2008. Language activism and the "new linguistics", Expanding opportunities for documenting endangered languages in Indonesia. In *Language Documentation and Description*, Vol 5, ed. by Peter K. Austin. London: School of Oriental and African Studies.

Francis, Norbert, and Pablo Rogelio Navarrete Gómez. 2009. Documentation and Language Learning: Separate Agendas or Complementary Tasks? *Language Documentation and Conservation* 3(2):176–191. Online: http://nflrc.hawaii.edu/ldc/http://hdl.handle.net/10125/4438.

von Gleich, Utta. 2005. Language documentation and development of teaching Material. *Language Archives Newsletter* 5:2–5.

Grinevald, Colette. 1998. Language endangerment in South America: a programmatic approach. In *Endangered Languages: Language Loss and Community Response*, ed. by Lenore A. Grenoble and Lindsay J. Whaley, 124–159. Cambridge: Cambridge University Press.

Grinevald, Colette. 2006. Worrying about ethics and wondering about informed consent: Fieldwork from an Americanist perspective. In *Lesser Known Languages in South Asia: Status and Policies, Case studies and Applications of Information Technology*, ed. by Anju Saxena and Lars Borin, 338–370. Berlin and New York: Mouton de Gruyter.

Hale, Kenneth L. 1992a. On Endangered languages and the safeguarding of diversity. *Language* 68(1):1–3.

Hale, Kenneth L. 1992b. Language endangerment and the human value of linguistic diversity. *Language* 68(1):35–42.

Hale, Sara Whitaker. 2001. Reminiscences of the trip to Australia 1959–1961. In *Forty Years on: Ken Hale and Australian Languages*, ed. by Jane Simpson, David Nash, Mary Laughren, Peter Austin and Barry Alpher, 19–28. Canberra: Pacific Linguistics.

Handman, Courtney. 2009. Language ideologies, endangered-language linguistics, and Christianization. *Language* 85(3):635–639.

Harrison, K. David. 2007. *When Languages Die. The Extinction of the World's Languages and the Erosion of Human Knowledge*. Oxford: Oxford University Press.

Hill, Jane H. 2002. "Expert rhetorics" in advocacy for endangered languages: Who Is listening, and what do they hear? *Journal of Linguistic Anthropology* 12(2):119–133.

Himmelmann, Nikolaus P. 2006. Language documentation: What is it and what is it good for? In *Essentials of Language Documentation (Trends in Linguistics. Studies and Monographs*, ed. by Jost Gippert, Nikolaus P. Himmelmann and Ulrike Mosel, 1–30. Berlin: Mouton de Gruyter.

Hinton, Leanne. 2003. Language revitalization. *Annual Review of Applied Linguistics* 23:44–57.

Holton, Gary. 2005. *Ethical practices in language documentation and archiving linguistic resources*. Presented at the Annual meeting of the Linguistic Society of America, Jan 6–9, 2005. Online: http://www.language-archives.org/events/olac05/olac-lsa05-holton.pdf

Holton, Gary. 2009. Relatively ethical: a comparison of linguistic research paradigms in Alaska and Indonesia. *Language Documentation and Conservation* 3(2):161–175.

Innes, Pamela. 2010. Ethical problems in archival research: Beyond accessibility. *Language and Communication* 30(3):198–203.

Krauss, Michael E. 2005. Athabascan Tone. In *Athabascan Prosody*, ed. by Sharon Hargus and Keren Rice, 55–136. (Current Issues in Linguistic Theory 269.) Amsterdam and Philadelphia: John Benjamins.

Labov, William. 1982. Objectivity and commitment in linguistic science: The case of the Black English trial in Ann Arbor. *Language in Society* 11:165–201.

Ladefoged, Peter. 1992. Another view of endangered languages. *Language* 68(4):809–11.

McNamara, Laura A. (2010). Moving forward with the CEAUSSIC Ethics Casebook: What is the Casebook, and why now? Blog of the American Anthropological Association, January 27. http://blog.aaanet.org/2010/01/27/ceaussic-ethics-casebook/

Mithun, Marianne. 2001. Who shapes the record: the speaker and the linguist. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 34–54. Cambridge: Cambridge University Press.

Montler, Timothy. 2007. The Klallam language program. UNESCO Register of Good Practices in Language Preservation. Online: http://www.unesco.org/culture/ich/index.php?pg=00145 &categ=05

Nathan, David. 2006. Thick interfaces: Mobilizing language documentation with multimedia. In *Essentials of language Documentation*, ed. by Jost Gippert, Nikolaus Himmelmann and Ulrike Mosel, 363–379. Berlin and New York: Mouton de Gruyter.

Newman, Paul, and Martha Ratliff. 2001. Introduction. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 1–14. Cambridge: Cambridge University Press.

Newman, Paul. 2007. Copyright essentials for linguists. *Language Documentation and Conservation* 1(1):28–43.

Olson, Kenneth S. 2009. SIL International: An emic view. *Language* 85(3):646–658.

Ostler, Nicholas, ed. 1998a. *Endangered Languages: What Role for the Specialist? Proceedings of the Second FEL Conference*. Bath: The Foundation for Endangered Languages.

Penfield, Susan D., Angelina Serratos, Benjamin V. Tucker, Amelia Flores, Gilford Harper, Johnny Hill Jr., and Nora Vasquez. 2008. Community collaborations: Best practices for North American Indigenous language documentation. *International Journal of the Sociology of Language* 191:187–202.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London and New York: Longman.

Raign, Kathryn Rosser. 2006. *The Decisive Writer*. Boston, MA: Thomson Wadsworth.

Research Center for Linguistic Typology. 2009. *Fieldwork Manual. Fieldwork and Your Wellbeing*. Bundoora: La Trobe University, Research Center for Linguistic Typology.

Rice, Keren. 2006. Ethical issues in linguistic fieldwork: an overview. Linguistic Fieldwork Preparation. A Guide for Linguists. Online: http://www.ch2006. ass.utoronto.ca/lingfieldwork/pdf/2.pdf.

Riley, Kathleen C. 2009. Who made the soup? Socializing the researcher and shaping her data. *Language and Communication*, 29:254–270.

Robinson, Laura C. 2010. Informed consent among analog people in a digital world. *Language and Communication*, 30(3):186–191.

Schneider, Cindy. 2009. *Why field linguists should pay more attention to applied linguistic research*. Paper presented at the Australian Linguistic Society Conference, Melbourne, July 2009.

Seeger, Anthony. 2001. Intellectual Property and Audiovisual Archives and Collections. *Folk Heritage Collections in Crisis*, May 2001, ed. by Council on Library and Information Resources. Washington D.C.: The American Folklife center, Library of Congress. Online: http://www.loc.gov/folklife/fhcc/propertykey.html.

Simpson, Jane, David Nash, Mary Laughren, Peter K. Austin and Barry Alpher, eds. 2001. *Forty* Y*ears on: Ken Hale and Australian languages.* (Pacific Linguistics 512). Canberra: Research School of Pacific and Asian Studies, The Australian National University.

Warner, Natasha, Quirina Luna, and Lynnika Butler. 2007. Ethics and Revitalization of Dormant languages: The Mutsun language. *Language Documentation and Conservation* 1(1):58–76. Online: http://nflrc.hawaii.edu/ldc/

Woodbury, Anthony C. 1993. A defense of the proposition: "When a language dies, culture dies." Proceedings of the First Annual Symposium About Language and Society-Austin (SALSA).) *Texas Linguistic Forum* 33:101–129.

Yamada, Racquel-María. 2007. Collaborative linguistic fieldwork: practical application of the empowerment model. Language Documentation and Conservation 1(2):257–282. Online: http://nflrc.hawaii.edu/ldc/

Chapter 7
Native Speakers and Fieldworkers

Fieldwork not only teaches you humility, it gives you friends for life. Abbi (2001:80)

7.1 Introduction

Having decided what language to work on and having arrived at the field site, the fieldworker must find native speakers to work with. Here are seven basic questions that s/he should consider when looking for speakers:

1. When should I begin looking for and hiring native speakers for fieldwork?
2. What is the role of each native speaker in my project?
3. What characteristics should the native speakers have?
4. How many speakers should I work with?
5. How do I maintain good relations with the native speakers I hire?
6. Do I have the needed permissions to work with native speakers?
7. What are my obligations towards speakers, and can I fulfill them?

7.2 Leads

Where can a fieldworker find native speakers? Many factors influence what the best leads are going to be. Abbi (2001:74–76) suggests getting leads on speakers from officials: good people to ask include schoolteachers, the postmaster, clergy, police, and resident missionaries. The problem with using officials as leads to finding consultants is that their position of power – think of the proverbially despised tax collector – may actually turn speakers off. They may not want to be associated with the official or may not want to give in to the official's suggestions that they participate in linguistic fieldwork. See Dixon (1984:32–33) for a good example of this. The fieldworker can follow several leads from several officials in order to keep an equal distance from each official (Abbi 2001:72). See also Paul's (1953) description of finding leads through official contacts quoted in Section 5.8.2.

S.L. Chelliah and W.J. de Reuse, *Handbook of Descriptive Linguistic Fieldwork*,
DOI 10.1007/978-90-481-9026-3_7, © Springer Science+Business Media B.V. 2011

It is very useful to find out if any speakers in the community have previously worked with a linguist or anthropologist. If other linguists have ties to a community, they might be helpful in getting the necessary permissions and finding speakers to work with. Ken Hale describes how, through the help of linguist Colette Grinevald and anthropologist Charlie Hale, he got permission to work on Ulwa and was assigned a speaker by the community (2001:77–78). Locating a consultant in this way did not give Hale a diverse speaker pool, but it made the fieldwork project more efficient.

Fieldworkers sometimes find speakers who are willing to interact with them by offering free literacy classes or English conversation classes. Fiona Mc Laughlin offered free English classes at her field site in Senegal in order to establish a network of contacts (Mc Laughlin and Sall 2001:193). In fact, since teaching and learning a language are easy concepts for non-linguists to understand, a good way to begin recruiting speakers is to announce the need for a language teacher (Paul 1953:432). A note of caution: some speech communities may not want an outsider to learn their language, so the linguist must have a feel for this before announcing his or her intentions to learn their language.

7.3 Timing

Judicious consideration of the factors discussed above will help in finding the most appropriate native speakers for the fieldwork project. Paul (1953:444) cautions that "…recognized deviants may be the first to offer their services." Speakers who first approach may not be "deviants", but they also may not be the most talented for fieldwork, so a fieldworker should bide his or her time in speaker selection. Dimmendaal (2001:62) suggests that one should never commit to long term work with any speaker until that speaker has been tried out. Rather, the fieldworker should hire speakers on a short term basis, withhold commitment for long term employment, and sift through the possible native speakers until it is determined who the appropriate speakers are. At some point a speaker may need to be dismissed because he or she is impossible to work with – for example, if that person is usually inebriated. It may also be necessary to reassign tasks according to speaker strengths. Crowley (2007:89) completely avoided entering into contractual agreements with speakers because firing is a lot harder than hiring! Instead, he used a much 'looser arrangement', where he visited speakers when their services are required.

7.4 Selection Based on the Role of the Native Speaker

When considering how to find and hire native speakers for linguistic fieldwork, it is useful to think of the specific role the speaker needs to fill. Will the native speaker be the primary consultant, expected to arrive for field sessions at a regularly scheduled time? Will the speaker have to answer specific questions to help with analysis of the phonetics, phonology, morphology, syntax, pragmatics, and semantics of the target language? Will the speaker be the primary source of data, or someone with whom previously elicited data is checked? Perhaps he or she will answer a questionnaire on their own time and share the results with the fieldworker. Is the speaker someone who is observed and recorded but not extensively interacted with? Perhaps the speaker is an "accidental" contributor to the fieldwork project, someone whom the fieldworker meets and works with for an hour, or meets at a social event and receives a mini grammar lesson from, or someone the fieldworker overheard at the market. Finally, one should know if the native speaker is a self-appointed or hired language teacher.

To get a perspective on these questions we first review the roles that native speakers have held in linguistic fieldwork in the past, and the implications that these roles have for the speaker/fieldworker relationship.

In linguistic and anthropological fieldwork conducted in colonial regimes, the native speaker was generally viewed as a peculiar object who aroused intellectual curiosity and therefore deserved study (Barnes 1963). Because native speakers were part of the colonized group, they did not have rights of privacy and had little power to hold back on the language and cultural knowledge they possessed. The fieldworker in turn did not have to justify how and why this data was being collected or how it would be used. Since the fieldworker assumed – rightly in most cases – that written documents resulting from fieldwork would not be accessible to the native speakers, s/he did not consider that his or her work would in any way affect the native speaker. Instead, descriptions of curious "primitive" peoples were expected to enrich the intellectual and cultural world of the colonizer. An extensive discussion of this point is given in Brumble (1983:285) who provides this illustrative quote from Kroeber.

> Why should we preserve Mohave values when they themselves cannot preserve them, and their descendants are likely to be indifferent? It is the future of our own world culture that these values can enrich, and our ultimate understandings grow wider as well as deeper thereby. (Kroeber 1900)

A consequence of viewing the native speaker as an object of study was that native-speaker contributions to anthropological and linguistic research were routinely ignored. Instead, fieldworkers cast themselves as the authoritative disseminators of native cultural and linguistic knowledge (Brumble 1983). Of course, the colonial outlook still continues today under other guises; for example, affluent, urbanized fieldworkers from a politically dominant group studying poorer, rural, politically subordinate communities. In a class and caste stratified society like India, it is easy for the educated and affluent to think of politically disempowered speakers as having little to offer other than primary linguistic data.

During and after the colonial period, linguistic fieldwork saw native speakers of little-known languages as people in need – in need of salvation through conversion to Christianity, and in need of the civilizing forces of Western society. Missionary linguistics – practiced by fieldworkers whose aim is to prepare religious materials for generating converts and who typically learn the language of the communities they work with – expanded the role of the native speaker from a subject of study to a teacher of language. Also, it is often the case that missionary linguists acknowledge speakers' contributions to fieldwork products such as Bible translations in native speaker biographies. An example is the biography of Patricio Xec Cuc who worked on the Quiché Bible described online at http://www.hbu.edu/hbu/Translations_.asp?SnID=2.

In the post-colonial era, there has been a gradual acknowledgement that research affects both the researched and the researcher (Cameron 1998). There are numerous examples of this. From a practical perspective, just the act of singling out and paying a speaker for linguistic work changes his or her position in the community. See Rice (2006:20) for a summary of the literature on this. Less obviously, as subjects of linguistic study become literate or interested in scholarship on their language, they see and read the material published about them, a fact that is especially possible with near universal Internet access. The image portrayed to the outside world of studied peoples and their languages is powerful in affecting both self and outside image. For example, the University of North Texas recently did a news story on endangered language documentation work conducted by us [Chelliah and de Reuse], and by our colleagues Tim Montler and Sadaf Munshi. Within days of this story being published on the internet, I [Chelliah] had comments, questions, and suggestions from Lamkang speakers expressing pride that their language was being recognized, and concern that the language be recognized correctly.

Because a native speaker's life is changed by research, today's field linguist recognizes the right of the speaker to have control over the direction of that change. See Rice (2006) for references and a review of the "empowerment" model of linguistic fieldwork. Thus today, the native speaker is more appropriately cast in the role of an individual with the ability and right to impose conditions and direction to linguistic fieldwork. The native speaker is empowered to be a co-constructor of the record and analyzer of his or her language (Cameron 1998; Cameron et al. 1992, 1993). The fieldworker facilitates this empowerment by recognizing it; the actual expression of this power – how much a particular speaker can shape the fieldwork – is determined by local politics and other factors such as the allocation of time and money to the project by the fieldworker. When I [Chelliah] began doing fieldwork in Manipur state on Manipuri in 1986, I remember that local linguists had easy access to speakers of Naga and Kuki-Chin languages. There was little concern about payment to consultants or the need to explain the purpose of linguistic research or to justify research products. By 2009, the situation had changed. I have been told by linguists in Manipur that speakers from those same communities are now taking a greater interest in the documentation of their linguistic heritage. They are also aware of the benefits that researchers gain from studying minority languages; they want to know what the products of the research are, and how the community will benefit from it. In my research on Lamkang, a language of southeast Manipur, I have found that younger speakers are concerned with the loss of knowledge of traditional stories and ritual language. They would like the documentation of these genres to be central to our linguistic fieldwork in the area. Thus for the Lamkang documentation project I am involved in, the speakers are the ones who shape the record because they determine who and what we record. See Mithun (2001:51) and England (1992) for more on this role of the native speaker.

We list here the labels used to refer to native speakers, each highlighting a different aspect of the speaker's role in the fieldwork setting. The role that a native speaker takes in a fieldwork project will change the output of that project. A particular speaker may also take on multiple or differing roles as fieldwork progresses.

- *Informant or Consultant*: Samarin (1967b:20) defines an informant as someone who provides samples of what has been said or what can be said in their language. Paul (1953:443) defines an *informant* as, "an articulate member of the studied culture who enters into a more or less personal relationship with the investigator for a relatively long period of time." As Udell (1972) points out, in a non-technical sense, the term *informant* has become synonymous with informer, one who rats on you'. Rice gives another reason for why *informant* is out of style: the title characterizes the native speaker as a "machine" that produces linguistic data (2006:25). For this reason, most fieldworkers now avoid *informant* and use *consultant* instead. However, Crowley (2007:86) says that *consultant*, too, can have a negative connotation in the developing world, where "self appointed experts are often contracted on highly paid short-term 'consultancies' to write reports that show little real awareness of the situation on the ground."
- *Subject*: Udell (1972) notes that Bloomfield used *subject*, a term which suggests that fieldwork is 'research on' rather than 'research with' a member of a speech community. Structuralists saw the main role of the native speaker as providing raw data, with the idea that this primary data alone was necessary for the fieldworker to arrive at a linguistic analysis. Native speaker input was also useful in gathering differential meaning, in which the native speaker compares two forms and explains the difference between them (Hoijer 1958:578). Since linguistic behavior was not considered to be modified by overall behavior, other input from speakers – such as how language should be used in daily situations – was not seen as necessary to analysis. The term *Subject* is also used in phonetic fieldwork when speakers participate in invasive phonetic experiments as described in Ladefoged (2003) and Maddieson (2001).
- *Native Speaker*: This term suggests a view of the speakers as a valuable sources of information, holding the key to the structure of the target language in their linguistic competence. However, it does not suggest speaker involvement in shaping the body of knowledge that results from fieldwork.
- *Field Assistant*: Crowley (2007:86) sees this as the role of the native speaker who, in addition to being a source of data, assists the linguist in making contacts within the community, sets up participatory research groups, negotiates commercial transactions, and/or helps with other project concerns.
- *Respondent*: In Hymesian ethnographic description and in sociolinguistic dialect or register studies, the term *respondent* is used for data gathered through interviews. Again, these interviews are controlled by and serve the research purposes of the interviewer.

- *Language Teacher*: The native speaker instructs the fieldworker in grammatically and socially acceptable verbal interaction in the target language, and teaches the fieldworker how to interact with other members of the community. The language teacher may also point out linguistic patterns which other second language learners find difficult to master (e.g. directionals, or terms of address and reference).
- *Research Assistant*: Typically, this term would be used for a native speaker who knows at least some linguistics and is able to help with either transcription or translation tasks, or who works with the researcher to analyze data and formulate generalizations. See Healey (1964) and Samarin (1967b:20–25). Udell (1972) uses *co-operator* or *co-worker* to describe this native speaker role. Nida (1981) and Crowley (2007:86) suggest *colleague* or *collaborator*.
- *Research co-investigator* or *partner*: Dixon (2010a:316) notes that we develop close relationships with our consultants, and build "priceless intellectual partnership[s]" with them. In Community-Based or Participatory (Action) language research, the native speaker takes on the role of co-investigator, bringing his or her own model of how to study language into play and sharing control of "establishing and creating knowledge about the language" (Czaykowska-Higgins 2009:41).
- *Co-author*: A speaker who writes or makes a substantial contribution to a publication. An example of this is linguist and native Manipuri speaker M.S. Ningomba who worked with linguist D.N.S Bhat in writing Bhat and Ningomba (1997), where the data are primarily from M.S. Ningomba.

The linguistic fieldworker, too, may take on different roles in a project. Below are some of the terms used to refer to fieldworkers, reflecting these varied roles.

- *Collector/Recorder/Analyzer*: The fieldworker records, elicits, and analyzes linguistic data using specialized training in linguistics and related fields.
- *Teacher*: In some cases – especially with endangered languages – speakers may come to think of the fieldworker as their language teacher. The native speaker learns some terminology and analytic techniques during field sessions and, if interested, s/he may also learn how to document their language and prepare pedagogical materials.
- *Student*: The fieldworker learns the language through at least some overt instruction from speakers. In these cases it can be useful to define the relationship between fieldworker and native speaker as student and teacher. However, this can become complicated if the native speaker decides that s/he must be the one to prepare and control elicitation sessions. Academic settings, such as Field Methods classes, are especially conducive to such roles.
- *Social Worker*: Consultants often ask for help or seem to need help with personal issues. Sadaf Munshi (p. c.) told us that one family wanted her to find a suitable boy for their daughter to marry. Fiona Mc Laughlin recounts the "moral crisis" as she "came to terms with the poverty" around her. She had to decide if she should use her money to solve the problems at her fieldwork site, or for her own research needs (McLaughlin and Sall 2001).
- *Mentor*: In some cases the native speaker is interested in studying Linguistics or following some other academic pursuit. The fieldworker may then proofread

or help write application letters, abstracts for conferences, or manuscripts for publication.

- *Friend, Spouse, Relative*: A new friendship can develop if the connection between speaker and fieldworker has been a long one. Existing relationships can turn into fieldworker–consultant relationships, for example when fieldwork is conducted with one's spouse or relative. The existing relationships will take precedence over and may interfere with the fieldwork relationship. At times, field communities may "adopt" the fieldworker. We [Chelliah and de Reuse] have been given clan affiliations with a language community in Manipur. Dixon was adopted as a "relative" in work in North Queensland, and was regarded as his consultant's mother's younger brother (Dixon 2010:317).
- *Project Manager*: Most fieldwork projects today are collaborative, and require team members with different areas of expertise. The field linguist may well find him/herself in the role of a project manager who must juggle budgets, payroll, and employee evaluations and re-hiring, in addition to linguistic and social concerns.
- *Employer*: Fieldworkers typically pay speakers for their time. If the relationship continues over a period of time, a fieldworker may become the sole source of income for the speaker. This is not an ideal situation, as the employment must inevitably end at some point. Dixon (2010a:315) says it is best not to look at the relationship with consultants as a business arrangement. It works best when the relationship is based on friendship.
- *Linguist*: Many Native American communities have a linguist permanently on staff at their cultural center, or otherwise affiliated in some capacity with their tribe. In these cases, the fieldworker can be referred to as 'our linguist' and native speakers can call themselves the language teachers (e.g., Timothy Montler's title with the Klallam nation (p.c.)).

There are some roles that a fieldworker should not take on: a fieldworker should not be the bully, cajoler, briber, or coercer of data from speakers. Individual personalities will, to a large extent, determine what role feels right to the fieldworker. For example, some linguists are shy people; they can be teachers, students and friends, but they may make terrible project managers. Or a linguist can be friendly, but not develop close friendships with people outside his or her own culture. For these reasons, a linguist cannot be considered a failure at fieldwork just because s/he cannot fill the same roles as a colleague. The fieldworker–native speaker relationship can be considered successful provided that "the relationship between fieldworker and the speakers of the language under study be an open and respectful one, and that the talents of the speakers be developed and put to use in a productive and creative way" (Newman and Ratliff 2001b:4).

It is clear, when reviewing the many different roles that the native speaker or fieldworker may take on in fieldwork situations, that no one title can exclusively characterize what each does. In this book, we choose *native speaker* as a cover term for these possibilities, as it is most neutral with respect to agency, i.e., a native speaker may be someone who contributes a narrative to the corpus, uses an interesting grammatical construction during an every-day interaction in the marketplace, helps or comments on linguistic analysis, or co-authors an article with the fieldworker. We use the term *consultant* when highlighting more active native-speaker roles in the fieldwork situation.

7.5 Selection Based on Speaker Characteristics

When selecting native speakers for fieldwork, fieldworkers consider specific characteristics, ranging from a speaker's intellectual curiosity to their physical condition. Ballmer (1981:63–64) lists the following prototypical native speaker features to consider[1]:

1. Producer: conceptualizes and articulates expressions of language
2. Analyzer: hears, understands, and processes expressions of a spoken language
3. Learning environment with respect to target language: primary socialization in target language
4. Age: adult, not a child or senile person
5. Health: receptors (eyes, ears, brain) and effectors (brain tongue, mouth) all operating normally
6. Intelligence: at least medium
7. Physical defects: should not exhibit a lisp or stutter
8. Linguistically proficient: can produce syntactically well-formed utterances that are semantically meaningful, pragmatically adequate, and textually coherent
9. Language competence: preferably monolingual
10. Linguistic environment: birth place, residence, and working place of the native speaker preferably in language community
11. Linguistic education: native speaker should preferably not be a writer, poet, literary critic, or linguist, but a normal language user
12. Performance abilities: a good communicator but not a professional performer
13. Intuitions: well-working
14. Introspection: does not need to offer grammatical judgments
15. Argumentation abilities: well-working
16. Theoretical abilities: not needed.

In this section, we review some of these features in detail.

7.5.1 Attitudes Towards Language

The native speaker who will serve as a consultant should exhibit a positive attitude towards the target language. As exemplified in Vaux et al. (2007:14–15),

---

[1] We have copied the wording from Ballmer's list, with some modification to fit the terminology used in this book.

if a language is socially stigmatized, speakers may not understand why someone is interested in it and may be reluctant to admit they speak it. Just because one speaker feels this way, it doesn't mean that everyone does. One should back off from the reticent speaker and look for the language enthusiast. A sign of a true enthusiast may be, as observed by Dorian (2001), someone who is critical of those who use anything but the target language.

Yamamoto (1998:220–221) discusses another aspect of this issue. The field-worker might find speakers who consider the ancestral language to be sacred and to be used only in ritual practices. Elders might feel that if the language can't be spoken well, then it shouldn't be spoken at all. In the case of endangered languages, this might mean a strong aversion to code mixing with a dominant language, e.g. some Native American languages and English in the Southwest. If the native speaker sees no place for the active use of the target language, and if they see English as needed for social advancement, then they will have little interest participating in linguistic fieldwork or learning more about their language so that it can be taught to others in the community.

7.5.2 Physical Condition and Age

The preferred age of a consultant is determined by specific fieldwork requirements. For phonetic and phonological fieldwork, the recommendations are as follows: Speakers should have all their own teeth and good hearing. Minimally a few upper incisors are needed to articulate plosives and fricatives clearly. There should be no speech impediment such as a lisp or stutter (Maddieson (2001) and Ladefoged (2003)). Maddieson suggests that school age children are easiest to work with, elders are the most difficult because they have less motor control over articulators, and the vibration of the vocal cords can be irregular. Crowley, for example, talks of working with elderly speakers who no longer have front teeth; with them it is hard to hear the difference between alveolar stops, nasals, and laterals (2007:91). Ladefoged, on the other hand, recommends working with older women or men because these speakers represent the best or most respected speech patterns of the community. Abbi (2001:77), too, notes that elders may be most suitable for phonetic work if they are the only ones who have retained a particular sound. In general, speakers should be healthy, and strong enough to concentrate on the tasks at hand for at least 2 hours at a stretch.

For other types of fieldwork, the usual recommendation is that speakers should be at least 16 years old, because younger children usually have not fully developed linguistically before that (see Healey 1964:345). Furthermore, special internal review board clearances are required for working with speakers under 18.

Nida (1949) points out that community elders often make the best consultants in terms of availability because they do not leave for work every day and so there is a better chance of finding them at home at predictable hours. In terms of data, Crowley (2007:90–91) warns that while older speakers can provide excellent narratives, they may not be as good in analytic tasks such as providing verb paradigms. Also, older speakers may be the most prescriptive and may provide archaic or formal speech samples, but not every-day language. They may easily tire during elicitation sessions and might not see the importance of paradigm elicitation, which they might consider to be repeated questioning over the same point. I [de Reuse] experienced something like this with some older Athabascan consultants who, as Crowley reported too, said, “I’ve already told you that word, why are you asking me again?” (2007:91).

In endangered language situations, one must work with the few speakers remaining. It may be necessary to work with someone who stutters. (Ivy Doak (p.c.) tells us that this made working on Salish reduplication quite interesting!). It may be necessary to work with someone in the early stages of Alzheimer’s, in which case their language skills may be excellent, but they may not be able to recall or repeat what they hear from a recording or even what they just said. We have experienced the same problem with chronic drug abusers; even if they are not ‘under the influence’ during a fieldwork session, the long term effects of drug use are obvious. Older speakers may be able to recall some lexical categories but may have trouble with recall of others.

There are always differences between the speech of older and younger speakers but, because of language shift and attrition, these differences may be more pronounced in some languages. In these cases, work with younger speakers must be supplemented with language samples from elders. Vaux et al. (2007) point out that, while younger speakers are more likely to defer to the fieldworker, they also may have short attention spans. This was certainly my [Chelliah’s] experience: 20-something males usually have something better to do than help transcribe a conversation. Excuses for not working ranged from, “I can’t come today, I’ve been invited to dinner,” to, “I can’t come today, I’m going to Agra to see the Taj Mahal with my cousin.” How can fieldwork compete with the Taj Mahal?

With older speakers it will often be necessary to have a handler who can act as an “interpreter” between the consultant and fieldworker. I [de Reuse] worked with one older speaker who was hard-of-hearing. His sister helped during our sessions by yelling elicited forms into his ear, using sign language to convey what was required in the target language, and letting me know when the speaker was too tired to continue with fieldwork.

7.5.3 Gender

In some areas of the world, it may be socially unacceptable to work with members of the opposite gender. The fieldworker may be completed barred from doing so in conservative societies. Dimmendaal (2001:60) notes that sometimes consultant spouses become jealous when the fieldworker is of the opposite sex. The fieldworker may have to prove that his or her intent is fieldwork and not romance. In Manipur, for example, social contact between a male and female often results in gossip which is sometimes malicious. When working with male speakers, I [Chelliah] left open the doors to the room we were working in, and sat with my consultant in front of open windows in full view of the street. This set-up satisfied everyone that nothing gossip-worthy was going on, and we were left in peace to continue with our work. Another possibility is to have a chaperon in the room at all times, keeping in mind that the chaperon might also expect payment. As discussed in Nida (1949), Healey (1964), and Bowern (2008:133–134), a fieldworker should find some way to show that he or she is sensitive to and respectful of local standards. This is a practical standpoint that the fieldworker must take in order to get work done. Also, one should not assume that all "conservative" societies are conservative in the same way. In many parts of Northeast India, it is acceptable for a woman to smoke, but it is not acceptable for a woman to talk directly to her brother-in-law. Conversely, in conservative urban areas, affluent women do not smoke, but taboos on inter-family conversation are not strongly observed.

Even given restrictions on working with the opposite gender, the fieldworker should definitely try to get data from both men and women, as this can expand the corpus in different ways. As Abbi (2001:76) points out, in traditional patriarchal societies, men have a wider network of acquaintances and experiences; thus, for a range of vocabulary and discourse data, men may make the best consultants. Since men are often involved in trade, they would also know the trade language well. If the trade language is being used as the contact language, this would also help in translation tasks (Nida 1949:190). How can the fieldworker respect native rules of cross-gender interaction while still gathering data from consultants? Foley (1991:132) suggests using male–female investigative teams for this. He reports that cultural restrictions against male and female interactions in New Guinea made it difficult for a male fieldworker to collect data from elderly women, the outstanding speakers of the community. In this case, a second person, a female native speaker, was hired to help make contacts and overcome this barrier to interaction. Bowern (2008: 134) suggests that if one is restricted from interacting with a speaker, a tape recorder can be given to an intermediary who can make recordings.

In terms of work dynamics, many fieldworkers prefer working with a speaker of the same gender because, "an informant of the same sex …will be more likely to feel comfortable interacting with the field worker, particularly when discussing potentially touchy issues such as terms for body parts, incontinence, and so on" (Vaux and Cooper 1999:7–8). On the other hand, we have heard of competitiveness creeping into fieldwork sessions when consultants were the same age and gender as the fieldworker. It is impossible to state a hard and fast rule in this area, since so much depends on cultural mores and the personality of the participants.

7.5.4 Location

It may be useful to begin studying a language by working with speakers in an urban environment. In India, for example, younger speakers of Tibeto–Burman languages can often be found in large cosmopolitan cities where they come to attend university, or work in call centers or in the Information Technology sector. But there are some challenges with engaging speakers in such environments. First, as Crowley

(2007:86–87) notes, the speaker certainly 'did not come to your city in order to become somebody's language–helper'. Speakers have difficulty in setting aside time in their day, or may have little patience for participating in elicitation sessions after a long day of work. The prospect of making a little bit of extra money will not necessarily be an incentive for speakers to show up for scheduled sessions. Crowley also notes that speakers who have lived away from their communities for an extended period of time in urban areas, in contact with other languages, and perhaps not often speaking their native language, may feel that they are not fluent enough to take on the role of a linguist's consultant. The fieldworker would then have to "reassure [the] reluctant language-helper that he or she really *can* help you in a meaningful way." (Crowley 2007:87). For example, they may just provide a basic word list and common phrases.

Not all speakers in urban areas will be happy to work with or even talk to a linguist. This might happen if, for example, the speakers and the segregated urban community they belong to are trying to keep a low profile in the larger community because they are not legal immigrants. It is never advisable to work with illegal immigrants in the United States because they cannot be paid without a valid social security number. Illegal immigrants will also be reluctant to sign IRB forms. Similarly, speakers from communities that have undergone the trauma of relocation, such as the Tibetan population in Delhi, may shy away from inquisitive linguists.

The speaker who is integrated into the larger community is perhaps more approachable. This type of speaker may associate work with pay and be a more dependable assistant. On the other hand, the integrated speaker may not want others to recognize that he or she speaks a minority language if it is stigmatized, and may not admit to speaking the target language. Crowley (2007:180) gives the example of indigenous peoples of Mexico and Central America who "try very hard not to be identified as Indians, especially when they are away from their communities," since being Indian or speaking an Indian language is stigmatized in these areas. For the most part, however, one is likely to get positive responses from speakers for requests to learn about their language (Crowley 2007:87).

7.5.5 Education and Literacy

It is very helpful to have at least one or two consultants who are literate in the contact language (Bouquiaux et al. 1992). If they can read and write in the target language as well, that is even better. These speakers can help with translations of previously recorded data. In some cases, such speakers can also work independently to provide word-for-word translations of transcribed texts. In Manipur, I [Chelliah] worked with consultants who had some linguistic training. Tasks that I gave them included checking and correcting transcriptions, filling in questionnaires that formed the basis for further elicitation, translating traditional grammars into English, and translating children's story books to English.

For phonetic fieldwork, a speaker who can write can be asked to jot down minimal pairs (Maddieson 2001:217).

There are disadvantages as well as advantages to working with educated speakers. As pointed out by Bowern (2008:132), if education is valued by a community, members would naturally understand or expect the more educated speakers to be included in an academic enterprise. However, educated speakers tend to be busier in that they typically hold office jobs with regular hours. In some cases, when educated speakers are hired for fieldwork, they are not interested in following the agenda of the researcher. For instance, Samarin (1967a:165) reports on work in the Central African Republic where literate speakers found fieldwork on ideophones uninteresting, while speakers with little education were willing to do the required work on this topic. Alec Coupe (p.c.) cautions that an educated speaker who has a position of authority in the community may feel empowered to construct explanations for cultural practices or terms encountered in texts that they do not, in fact, genuinely understand.

One should avoid using *only* educated consultants, since these speakers may only provide prescriptively correct forms, consider their idiolect to be linguistically prestigious, and/or be critical of stigmatized variants in other dialects (Kibrik 1977). Nancy Caplow (p.c.) reports that in the specific case of Tibetan, Sprigg (1991) has pointed out that a hazard of working with educated speakers is that their pronunciation of elicited forms can be dramatically influenced by their knowledge of Written Tibetan. In fact, he finds that three styles of pronunciation can be distinguished: “Ordinary”, reading, and spelling.

Of course, prescriptivism can be a factor whether or not a speaker is formally educated or literate, since speakers often have an opinion regarding what is “better” or “more correct” speech, based on non-linguistic prestige and stigma, determined by a variety of social factors.

Language teachers can be good consultants who can provide eloquent grammatical explanations (Abbi 2001:75–76). Of course, the fieldworker must be careful about the prescriptivism in this case as well. Abbi suggests working with teachers after getting a feel for the main structures of the target language.

Needless to say, a speaker need not be educated to be a good consultant (Dimmendaal 2001:64–65). Uneducated speakers may, in fact, know more about the oral history of a community simply because they have spent more time in the community and less time at school (Bowern 2008:131).

7.5.6 Personality Traits

The fieldworker must consider a potential consultant’s personality traits. Fieldwork requires a consultant who is patient and does not mind repeating the same form several times. The consultant should be reliable and show up on time. A common fieldwork frustration, as reflected in Macaulay’s (2004) experiences, comes from scheduling an appointment, preparing for a session, and then being stood up. In fact, the fieldworker should expect to be stood up. It is unreasonable to expect every single speaker to show up as scheduled. The fieldworker should consider what it would be like if someone showed up in his or her neighborhood and requested help with their research at such and such a time, several days a week. It would be difficult to rearrange one's life to suit the needs of this stranger. Whatever time a speaker can spare is a gift, even if they are paid for their time. As discussed in Bowern (2008:135–136), to keep sane in the field, the fieldworker might rank consultants by their reliability. Less reliable consultants can be employed for more flexible work, such as double-checking data already analyzed, checking transcriptions, or being recorded having conversations with others.

To some extent, the personality of the speaker should match that of the field-worker. Speakers with difficult personalities, or someone who argues or is easily offended, will not make a good primary consultant. Consultants should be outgoing. Nida (1949:190) states that, "the reticent, taciturn person rarely qualifies as an informant". Speakers should have confidence in their knowledge of the language, even though they have not studied the language in an academic setting. Mithun (2001) tell us that speakers are sometimes unwilling to claim that they know the language well because it was not learned or studied at school.

Perhaps the most important trait to look for in a speaker is mental alertness and intellectual curiosity (Nida 1949:190; Dimmendaal 2001:61). Speakers who are interested primarily in checking up on the fieldworker's activities and gossiping are not going to make good primary consultants – but they can still be helpful in learning about conversational interaction. Speakers with limited attention spans and who bore easily will frustrate the fieldworker. Their responses to session questions may be unreliable because their main goal might be to end the session or the fieldworker's line of inquiry. One of my [Chelliah] consultants would consistently reply "maybe" to Yes–No questions when he was tired. This was a sign that it was time to end our session or move to another line of questioning. See also Samarin (1967b:140).

7.5.7 Talent as Consultant

In addition to the general personality traits discussed above, there are intellectual and linguistic abilities which make a native speaker a good consultant.

- *Have Language Skills*: Here, we do not refer to fluency or native proficiency; there is more on those points below. Rather, we mean that the speaker should have language-based skills, e.g. pronounce clearly (Kibrik 1977); be a good storyteller (Everett 2001); be interested in talking about word histories and meanings; and have some interest in language and how it connects with the history and culture of the speech community (Abbi 2001:74).
- *Provides Extended Responses*: Speakers should be able to give extended responses to requests for specific forms. For example, if the fieldworker asks for a particular form, a good consultant might provide a story or hypothetical situation, and within that, a clause where the requested form is used. To be successful

at this, consultants should not interpret questions too narrowly. See discussion in Hopkins and Furbee (1991:65) and Kibrik (1977).

- *Provides Useful Explanations*: Some speakers understand the purpose of linguistic fieldwork and are able to talk about the relationship between form and meaning (Abbi 2001:79). Dixon (1992) provides examples of how a speaker might relay grammatical information without linguistic terminology. The speaker may not be able to explain word class membership using terms such as 'verb' and 'noun'. However, they may relate the difference by providing paradigms so that the fieldworker could identify a word as a verb because it conjugates like a verb. Or, a speaker might illustrate synonymy in the target language by giving an example of synonyms in the contact language. Speakers will have trouble describing the meaning of tense or aspect morphemes (Matthewson 2004:384); even so, talented speakers will find a way to get the message across. See Kibrik (1977) and Dixon (1992) for further discussion on the linguistic sophistication of consultants.
- *Is Enthusiastic About the Fieldwork Project*: The preferred consultant shows initiative in adding to data and analysis. For example, he or she might double-check facts discussed during elicitation sessions with other members of the community (Dimmendaal 2001:63). The enthusiastic consultant sticks with the task at hand until it is completed, and is not frustrated when fieldwork progresses slowly (Healey 1964). Finally, the consultant shows a concern for accuracy by correcting the investigator's mistakes (Healey 1964:345). The fieldworker should be careful about speakers who are too willing to agree with all the comments or attempted target-language utterances produced by the fieldworker. Also, if a speaker is unwilling to provide any information on the language, then the fieldworker should find other speakers to work with (Abbi 2001:78).
- *Is Able to Repeat Forms*: Repetition is almost a form of analysis, since, in repetition, the communicative import of an utterance must be separated from the actual structure of the utterance. Some speakers simply cannot repeat a word out of context, and therefore cannot repeat a word the same way twice. Other speakers can repeat an utterance or part of an utterance several times in exactly the same way, and this helps the fieldworker transcribe the utterance accurately. If a speaker is not a good repeater, lexical tone might be different in repeated tokens. Finally, some speakers will 'regularize' or 'correct' the utterance – for example, by changing fast speech forms to careful speech, or by reverting to a grammatically prescriptive norm.
- *Is Able to Translate Accurately*: Some speakers, even if fluent in the contact and target languages, may find it hard to translate from one language to the other. The preferred linguistic consultant is one who can provide a translation that captures both the denotative and connotative meanings of the target and contact languages. This is a specialized skill that all speakers cannot be expected to have. Sometimes it is necessary to consult a number of speakers before arriving at a satisfactory translation. In general, we define "satisfactory" as a translation that is accepted by two or more speakers. Beware of speakers who can only provide calque translations.

- *Is Trainable*: The best speakers to work with for an extended period are those who are open to training. See the extensive list in Healey (1964:348–349) of training tasks that can be set for the speaker-in-training. Working with a consultant who has already done some fieldwork can save quite a bit of time, as these speakers are already 'trained' to the elicitation tasks used. Scollon (1979) writes that training speakers for a linguistic project may not be an option in cultures where speakers resist creating a central knowledge repository; that is, since each person is considered an individual, the idea of a central bank of knowledge is unrealistic. For instance, Scollon (1979:40–41) observes that in Chipewyan (an Athabascan language of Canada) there is a type of "bush consciousness" that is opposed to recognizing or displaying the specialized knowledge that is useful for linguistic description. Thus trying to train a speaker becomes difficult because this same consciousness triggers a kind of resistance to any training or helping the researcher achieve his or her linguistic goals. In these cases, the fieldworker must rely more on data collection as a participant–observer, and on learning the language him/herself (Scollon 1979:40).

7.5.8 *Language Proficiency*

It seems obvious that linguistic consultants should be native and fluent speakers in the target language. The definitions of 'native' and 'fluent', however, are not equivalent across fieldwork situations. In particular, the notion of 'native speaker' is not universally accepted, as shown in Paikeday (1985).

In endangered-language speech communities, it is best to work with speakers whom the community regards as good speakers. Yamamoto (1998:222) explains that what constitutes a good speaker may have nothing to do with fluency, but may instead be based on extra-linguistic features such as clan membership, religious background, blood quantum, or knowledge of ritual language. Alec Coupe (p.c.) relates that one of his outstanding consultants was considered an inappropriate consultant by some members of the speech community because the consultant's father was not a member of the tribe.

When different varieties of language are used for ritual speech and colloquial speech, some speakers control the every-day variety, while others are actually non-speakers who have memorized ritual speech events; e.g. singers who memorize Chiwere (Siouan) songs in order to perform them, but who otherwise do not speak the language (Hopkins and Furbee 1991:65).

Ironically, as Yamamoto (1998: 221) also points out, non-fluent speakers may be the most eager to be involved in linguistic work with pedagogical aims. Furthermore, perfectly good speakers can sometimes be blackballed for reasons other than their linguistic ability. The fieldworker is more or less expected in some communities to work, for at least part of the time, with whomever is considered the 'good speaker'.

Evans (2001:258) reviews the labels 'rememberers', 'semi-speakers', and 'speakers' used in endangered language situations to label different fluency levels. Bowern

(2008:137–139) uses the terms ‘full versus part’ speaker and ‘semi or passive’ speaker. ‘Rememberers’ are speakers who can recall words and phrases, but who do not use the language on a day-to-day basis or speak it fluently. Their pronunciation may be strongly influenced by a “large” language. A ‘semi-speaker’ may have fluency in some registers, but may have limited mastery over some grammatical constructions and cannot recall or has never learned many lexical items. A ‘speaker’ has full communicative proficiency. Having community members accurately identify who falls at the ‘speaker’ or the ‘rememberer’ ends of the continuum can be problematic. Evans (2001:253–261) discusses a distinction observed in Australian communities between speakers who “own” a language and those who “speak” the language. Language ownership is conferred on a speaker “by descent based membership of particular social groups, such as clans.” (Evans 2001:254). A language owner may or may not be a fluent speaker of the language he or she owns; many speakers have an affiliation with one clan, and so “own” that language, but live with a different clan and thus actually speak a different language on a daily basis. In this case, the fieldworker must look specifically for speakers of the target language, and not simply for people affiliated with a clan that speaks the language. For political reasons, a speaker may act fluent but only have mastery over a very limited repertoire: for example, a clan leader might try to enhance his position by trying to display linguistic skill in the clan language; a speaker who belongs to a clan only peripherally might try to affiliate with the inner circle by using the corresponding language; or a speaker may exaggerate fluency to gain status in his own or another community (256–257). In all this, the more fluent speakers may be missed by the fieldworker. Although it is prudent to never discount a speaker who seems disfluent, the fieldworker should be aware that he or she might, because of these practices, miss the fluent speakers.

Fieldworkers offer several suggestions on how to test for speaker fluency. Grimes (1995) discusses how a test, originally designed by the US Foreign Service to score second language proficiency, can be adapted to assess fluency in a first language. A fieldworker might run this test at an early stage of speaker selection. The test assesses a speaker on aspects of linguistic behavior such as comprehension, discourse competence, structural precision, and lexicalization (appropriate choice of words).

Whaley and Li (2006) outline a seven part elicitation protocol for a “quick assessment of dialect and degree of fluency.” The protocol includes 20 common vocabulary items, paradigms for a verb in two tenses, sample causative and intransitive clauses, and words containing predictable derivational morphemes. Whaley and Li note that even though it does take several months to develop, there are clear advantages to preparing such a protocol, as listed here:

- Fixed elicitation protocol is ready to use when new speakers are encountered.
- The elicitation procedure usually takes under 1/2 an hour, but allows for reasonably accurate assessment of fluency and dialect.
- The process of elicitation helps speakers who haven’t used… [the target language] for a while ease back into thinking about it.
- The time spent eliciting … [the protocol] data also serves to identify speakers who may be particularly gifted as consultants.
- Allows for simple comparison of age groups and genders on the same tasks.

(Whaley and Li 2006)

Finally, when the fieldworker attempts to learn a language, he or she can assess which speakers are able to communicate freely in the target language. See Dimmendaal (2001).

When considering fluency, it is important to take extended networks into consideration. Linguistic performance is never homogenous throughout an entire speech community, since social networks can cause variation in the skills of individual speakers (Milroy 1980, 1987). Thus when judging the linguistic skill of a speaker, it is necessary to look at the strength of their family relationships, personal histories, and community interactions. Each of these factors determines language fluency and linguistic variation. (As shown in Kroskrity's 1993 study of English and Tewa discussed in the paragraph below, it is not always easy to predict how these networks influence fluency.)

Another factor relevant to the evaluation of speaker fluency is the nature of speaker bilingualism or multilingualism. In many cases, speakers mix freely between two or more languages. When listening to Lamkang speakers, I [Chelliah] hear them freely code switch and code mix between Lamkang, Manipuri, Hindi, and English. When distinct languages are used for different social situations, it is not clear if speakers are, in fact, proficient in all the languages they use. In these cases, it can be difficult to say if a speaker is fluent, or to decide what counts as fluency in the target language. Furthermore, as in all multilingual scenarios, the fieldworker must be careful about language contact influence. Scollon (1979:3–5), for example, describes a situation with the Chipewyan at Fort Chipewyan who speak Cree, Chipewyan, English, and French, where the phonology of these four languages is converging.

Perfectly balanced knowledge of two languages is rare. To know which language a speaker is most fluent in, one must know the function of that language in daily life. Additionally, as illustrated in Kroskrity's (1993) description of three Tewa speakers, it is useful to know a person's linguistic history. At the time of Kroskrity's writing, two speakers had "structural and use-oriented fluency" in Tewa and English. Their fluency in Tewa could be explained through strong positive ties to the community and daily exposure to older speakers. A third speaker grew up speaking Tewa in a conservative Tewa household, but he had weaker family and community ties, encouraged by a life lived on the periphery of the reservation. He spoke mostly English. Although his Tewa accent was still strong in terms of semantics, syntax, and morphology, this speaker's Tewa showed convergence with English patterns. His conservative upbringing did not ensure fluency in Tewa for this speaker, any more than the bilingualism of the two fluent speakers entailed disfluency in Tewa.

It is never a good idea to discount speakers' contributions on the basis of a perceived or even tested lack of fluency. Bowern warns against making speakers feel ashamed about not speaking fluently (2008:137). Semi-speakers may not possess all the qualities a linguist needs in a consultant, but they often possess some useable skills; e.g., perhaps they cannot produce narratives, but they can still translate them. Another important point Bowern makes here is that semi-speakers often become more fluent as they are given time to think in and about their language. Group discussions can help with reviving language memories as well.

A native speaker should not be dismissed just because others say that s/he is not fluent. Even in non-literate societies that do not have a codified standard, distinctions are made between speakers who are eloquent, too archaic, too preachy, illiterate, or stupid. Bloomfield (1964:396) states that speakers decide on what ‘good’ language is by considering “an accumulation of obvious superiorities, both of character and standing, as well as of language [because] some persons are felt to be better models of conduct and speech than others. Therefore, even in matters where the preference is not obvious, the forms that these same people use are felt to have the better flavor.”

In any case, it is useful for fieldworkers to note all data, whether it comes from speakers who are perceived to be fluent or not, because comparisons between speakers can contribute to studies of age-graded change or language decay. See, for example, Dorian’s report on her fieldwork on Gaelic (2001:143).

For a definition and discussion of the concept of ‘native speaker’, see Coulmas (1981), Paikeday (1985), and Davies (2003). A native speaker is not always the best choice for a field consultant because even though the speaker may have learned the target language in childhood, that variety might become underused and forgotten by adulthood (Davies 1994:2722). Second-language or heritage learners of the target language should not be discounted as consultants, because even if they cannot make reliable grammaticality judgments, they can still contribute data such as lexical items. Furthermore, second language learners may be better suited for linguistic fieldwork than some available native speakers because they excel in paraphrasing, hedging, and predicting or empathizing in conversation (Davies 1994:2720). It should be kept in mind that non-native speakers may need more background context for interpreting linguistic cues than native speakers (Davies 1994:2714).

7.5.9 Availability

Primary consultants will be speakers who are willing to set aside at least 2–3 hours at a stretch for linguistic fieldwork. Sometimes even the most dependable speakers will disappear whenever a festival, political campaign, wedding, or other social event takes place. This is to be expected; the fieldworker cannot assume that the fieldwork session will be a priority for the speaker, especially in the first few months when the fieldworker may not yet be seen as significant to the community. See also Dimmendaal (2001:60) and Bowern (2008: 135–136) on the topic of scheduling and availability.

7.5.10 Objectives

When selecting consultants for a linguistic project, the fieldworker should select speakers whose objectives do not contradict his or her own. For example, a speaker may mainly be interested in being seen with the fieldworker to show that he or she has influential friends, or in making money. Neither of these goals stands in the way of fieldwork objectives. Equally important, the objectives of the fieldworker should be aligned to the objectives of the speaker. Okuda (1995) reports on some Ainu speakers who wanted to document conversational interaction, while others were interested in preserving the language of traditional rituals. In this case, an investigator wanting to document Ainu grammar would not be well-received by either group unless each group felt that its goals were going to be met. See Samarin (1967b: 20–25) for further discussion.

7.6 Number of Consultants

For phonetic fieldwork, the recommendation is to record multiple speakers, both male and female. See Maddieson (2001:221), for example. For other areas of investigation, fieldworkers recommend anywhere from one to as many speakers as possible. This variation is determined by practical considerations and by the fieldworker's beliefs as to what extent an individual's linguistic competence provides a representative grammar of a language.

7.6.1 Linguistic Competence and Language Use

A fieldworker who believes that only individual grammars exist will be satisfied with data from a single speaker. To Vaux et al. (2007:8–9), studying a single speaker provides consistent data, while, according to them, working with two or more speakers may be "counterproductive" because this will introduce potentially confusing idiolect and dialect variation. The same idea is argued for in Wong (1975:45), but for different reasons. He arrives at "the disturbing conclusion that the majority of the [descriptive] grammars which we have at hand are no more than conglomerations of individual [mental] grammars, or grammars of the people who write them, but not [mental] grammars of the entire language they purport to describe." In this view, descriptions of single competencies should be just as interesting and perhaps even more accurately representative of an existing grammatical system than descriptions based on input from several speakers.

While the reasoning for working with one speaker is theoretically attractive, it is not justifiable. In order for a theory in the hard and social sciences to be considered quantitatively rigorous and adequately statistically reliable and valid, the researcher must show that s/he has controlled for variability among subjects, display trends or occurrences based on a statistically significant number of respondents or responses, have results that are replicable, and use methodology that controls for data skewing (Larson-Hall 2009; Mackey and Gass 2005). We advocate holding the descriptions that result from linguistic fieldwork to the same standards. For the results to be valid, data collection should follow the same rigorous standards placed on any scientific investigation. Data collection methods should provide data that are robust, replicable, and quantifiable. The evidence for analysis should be independently verifiable (Mosel 2006:53).

Working with more than one consultant can go a far distance towards meeting those standards. Data from varied sources can guard against distortions resulting from dressage, the observer's paradox, faulty questioning, or prescriptive influences of one individual's idiolect. Working with several speakers will provide the researcher with points of comparison so that he or she can learn to distinguish between reliable and unreliable data.

It is true that when one gets language data from a number of speakers, intra-speaker variation will be observed. It is in this variation that we can learn what constructions are normative and part of a shared linguistic system (Bernard et al. 1984:512). Variation – whether determined by register, style, genre, interlocutor, gender, or socio-economic level – provides clues about linguistic structure. See Bloomfield (1964:394–395) and Labov (1969).

In relatively robust Australian aboriginal communities, it is likely that speakers will try to keep dialects distinct (Sutton and Walsh 1979:23). But dialect and language mixing can occur, and when it does, it can pose a serious problem for language description if only one speaker is consulted. As Gerritsen (2004:84–99) concludes from study of an Australian language, unless the fieldworker is fully aware that mixing is occurring, and can identify the languages which are being mixed, a historical appreciation of the language as well as an understanding of the linguistic system of the target language will be flawed.

One should always take all speakers into consideration when doing "salvage" linguistics. When a language is robust, one can always collect data and then check it at a later time with other speakers. This is not the case where there are only a few speakers remaining. The fieldworker needs as much native speaker input as possible; for endangered languages, the time to get that input is now. So, as Dixon did, it might be necessary to use "detective work" to find all possible speakers (Dixon 1984).

Carden (1990) writes about the possibility that "randomly distributed dialects" exist. These are dialects that do not correspond with a particular geographic area or social class; they are oddball dialects that originate due to particular social histories. Thus the fieldworker cannot rely on a single speaker for data. Grammaticality judgments from a single speaker are also of questionable usefulness; data from other speakers is required.

There can also be also variability in a speaker's responses between sessions. This may be due to fatigue, the elicitation method used, boredom, or other extra-linguistic factors. On this point, see Grimes (1995:10) for discussion of word list elicitation.

7.6.2 Practical Considerations

There are practical considerations that should lead the fieldworker to work with more than one speaker.

- *Economics*: Working with more than one speaker will ensure that the fieldworker is not expected to be the sole economic support for that speaker. It is also important that fieldworkers be seen as spreading the resources to the community and not privileging one person over others. The smaller the community, the more important this will be.

- *Varied Talents*: Speakers have varied talents, so using just one speaker would limit the corpus. In addition to the speaker characteristics given in Section 7.5, consider the following specialized knowledge of abilities discussed by Dimmendaal (2001) and Mithun (2001): specialized vocabulary; social significance of grammatical constructions; ability to provide paradigms; interest in etymology or compositional semantics; and sensitivity to linguistic detail.
- *Speaker Fatigue*: Reflecting on language is not normal everyday behavior. It is difficult and tiring. Therefore, it is important to work with more than one person, and that, too, each for short periods of time. Recommendations on how many sessions to have per day differ. Crowley (2007) recommends one session per day, with the rest of the day reserved for review and for preparation for the next day. The fieldworker might want to have at least two speakers lined up per day since (1) a consultant may tire after 2 or 3 hours of work; (2) one of the consultants might not show up. Healey (1964:345) recommends working with a different speaker each day of the week. If the fieldworker plans a different activity for each day based on the skill of the scheduled speaker for the day, no one gets bored and the fieldworker has plenty of time to process information between sessions. Healey also advises against using the same person for household help and linguistic work because there will be too many demands on the speaker's time, and this will result in speaker fatigue and will make scheduling sessions a problem (1975:346).

7.6.3 Group Dynamics

It is also possible to work with two or more consultants at the same time. When setting up such group sessions, general issues of group work and group dynamics come into play. When speakers are in groups, personal animosity, shyness, or embarrassment about degree of fluency may prevent them from conversing even if they are the only remaining speakers of the language under study. When speakers are very close in age or friends in daily life, they may compete against each other in fieldwork sessions, in a sense "showing off" for the fieldworker and one-upping their friend. In some Australian aboriginal communities, members of a dyad may be related to each other such that conversation between them is either "difficult or entirely prohibited" (see Healey 1964:346). In mixed gender groups, male speakers may try to control the female consultant's responses during fieldwork sessions. This is especially true, in our experience, with married couples, where one person usually takes the lead role. In general, it seems to us that the more consultants there are participating in a single session, the harder it is to get the most out each individual's talents and to keep everyone engaged and feeling useful.

On the other hand, there are certain combinations of consultants that work very well: I [Chelliah] have worked with a teacher/student pair and a father/son pair. The relationship between the members of these dyads was such that they could freely discuss the data without being a threat to one another, and each had a different useful piece of information to contribute to the question at hand.

7.7 Collective Fieldwork

Thus far we have talked about fieldwork as if it is always conducted by a lone fieldworker, working for extended periods of time in a community. This is the dominant linguistic fieldwork paradigm as described for the Leningrad school of fieldwork in Aikhenvald (2007), and described also in Samarin (1967b), Crowley (2007), Bowern (2008), and in the articles in Newman and Ratliff (2001a). A different approach to linguistic fieldwork, a collective approach that uses the concept of division of labor, is described in Kibrik (2007). Originally developed at the Lomonosov State University of Moscow for the training of undergraduate students in fieldwork, this method is now used as a way of efficiently documenting little known languages. As described in Kibrik (2007:38), collective fieldwork is carried out by hierarchically organized groups: a Head researcher, faculty members and experienced fieldworkers, and beginning students. We provide part of the description of the field-trip for documenting Khinalug (Northeast Caucasian) in Kibrik (2007:31) as an example of collective fieldwork:

> The grammatical description of Khinalug is based on data gathered by a team of fourteen people in the summer of 1970. The team consisted of 3 researchers and 11 students. Each researcher was personally responsible for one of the following three domains: phonetics…, lexicon and texts…, and morphology. The students collecting morphological data were split into groups and each group was in charge of a particular topic in morphology.

In this approach, days are spent in data collection. In evening "workshops", teams working on the same topic meet to share information and develop preliminary analyses (Kibrik 2007:40–41). All activities are controlled by the head, who is responsible for the design of the field-trip, including practical matters such as lodging, helping to bring about a consensus view when there are competing analyses, making theoretical decisions for the group, and hearing reports from group leaders (Kibrik 2007:38).

The field-trips are typically short; for instance, the trip to study Khinalug lasted 6 weeks. Field-trips are typically followed by analysis and grammar writing conducted away from the field by the more experienced linguists. This is followed by another field-trip for data and analysis checking.

There are several advantages to the collective fieldwork approach. For one thing, different expertise can improve and accelerate data collection and analysis; for example, when one researcher is weak on morphological analysis, another can help in that area. Also, a large amount of data, in the form of texts, wordlists and questionnaires, can be collected in a short amount of time while consultants are available. Because fieldwork is conducted quickly and resources are shared, costs are not prohibitive.

There are also, unfortunately, some disadvantages to collective fieldwork. There is often an uneven ratio between consultants and fieldworkers, so that fieldworkers must compete to get a consultant or schedule time with the best consultants (Andrei Kibrik and Anastasia Bonch-Osmolovskaya p.c.). Consultants may not be available or willing to work when the team arrives at the field site. A single researcher can reschedule or move to a different site to find other speakers, but a larger group will not have this scheduling flexibility. The short stay also prevents development of a deep understanding of the culture, and more or less ensures that fieldworkers will not learn the target language. Finally, whatever mechanism is in place for correctly attributing a fieldworker's contribution to a project may not always adequately match a fieldworker's expectations. In this system, a fieldworker cannot claim as his or her private intellectual property an analysis of data, and s/he must be ready to accept that. As Kibrik (2007:30) puts it teamwork "excludes manifestations of absolute individualism, and therefore is not suited to all researchers."

7.8 Group Participatory Research

The goal of group participatory research is to empower speakers to document their own language for the purposes that they deem necessary. An example of this is the model of research described in Benedicto and McLean (2002), where Mayangna women of Nicaragua were trained to do basic documentation work such as recording, transcribing, preparing dictionary entries, and preparing texts for publication in the standard orthography. I [Chelliah] have been involved in a similar project with the Lamkang of Manipur State, India. Our team works in the following way: in addition to my work with community members, four Lamkang speakers record speech events that they feel should be documented (mainly traditional folktales, stories about clan origins, and folk songs). With the help of a trained Manipuri linguist, the recordings are transferred to our project computer and then transcribed in the Lamkang orthography and aligned to the speech signal. When I meet with the speakers and the Manipuri linguist, we check the transcriptions and complete text translations. Participatory linguistic fieldwork increases the number and types of consultant involved in the project. No participant is just a consultant, and no one is strictly the only researcher. Instead the roles are shared, as are the goals and outcome of the research.

Dwyer (2006:55–56) provides the following practical steps for setting up a team. The wording in quotations is directly from Dwyer. The quotation is followed by a paraphrase and expansion of her explanation.

1. "Assemble a team of local colleagues": At the pilot-study stage, visit the field site, make contacts, and identify potential team members.
2. "Propose a research plan": By talking to interested speakers, discover what their needs and interests are for language documentation or pedagogy. Use this information to apply for funding.

3. "Narrow the scope consultatively": Make a list of speech events the community wants documented. When investigating a particular aspect of grammar, such as tone, the fieldworker must identify which speakers to work with. If the fieldworker intends to administer a sociolinguistic survey, team members should be trained so that they can administer the survey and catalog the results, and know how to appropriately inform the community of the upcoming research activity.
4. "Archive materials locally and remotely": Make arrangements to have the materials archived locally. Teach team members how to access and use online archives.
5. "Work with small, stable, offline software" and "Work with computer programs with which your local partners are comfortable": Set up computers with the necessary software. Use free software for transcription and analysis such as TRANSCRIBER or TOOLBOX. Write easy-to-follow project-specific manuals for team members. Review and demonstrate recording, transcription, and translation procedures several times.
6. "Keep checking in with team members": Check regularly to see if there are issues with software or hardware. Encourage team members by reminding them of the larger goals, and of how much has been done and how much needs to be done. If payments are involved, keep payments flowing at regular intervals to encourage consistent work.
7. "Make sure the local researchers see interim and final products": Share as many products of collaborative work as possible. Always share something at regular intervals to keep enthusiasm for the project going and to build and maintain trust between the fieldworker and the speakers.

Setting up and running a group project takes management skill. It requires the fieldworker to be in charge of everything, from payroll to personnel issues, and to coordinate work-flow for the project and for all the project members. It requires development of training materials for participant speakers, and mentorship of graduate students who are working on the project. Parallel to the steps for documentation set out by Dwyer are steps for program management discussed in Glenn (2009). These are:

- Coordination: manage participants, centralize and synthesize input, keep research goals clear
- Distribution of Labor: manage virtual designated workspaces, define participant roles
- Interoperability: ensure portability and standard presentation and coding of data
- Authorship and Authority: manage issues of intellectual property and copyright
- Feedback: enable continuous evaluation of project products

As stated by Czaykowska-Higgins (2009:43), the linguist's role should never be that of "a knight in shining armor who rides into a community to rescue it." The collaborative model allows for balance between the roles of speakers and linguists in the description and documentation of a language.

7.9 Long-Distance Consultants

In some situations it is possible to train consultants so they can complete aspects of the fieldwork on their own. One example where this was successfully carried out is described by Hale (2001:97) for the compilation of the Ulwa dictionary. Equivalent Ulwa forms from an existing database of another Nicaraguan language, Miskitu, were collected. Sentences in Ulwa were also collected to illustrate word usage. Once Hale left the field, this work could be continued, because consultants who could write in the Miskitu orthography wrote down the Ulwa utterances using the Miskitu orthography. In such instances, the fieldworker would need to check all recorded forms at a later date and transcribe them phonetically, since native speakers have a hard time hearing phonetic distinctions; even if these are pointed out during training, they can easily slip from attention (Hoijer 1958: 581).

7.10 Linguist Native Speaker

Ehlich (1981:154) distinguishes between what he calls the “native native-speaker” and the “linguist native-speaker”. The native native-speaker is simply a native speaker of the language who is not trained as a linguist. The linguist native-speaker is a trained linguist who is carrying out documentation and description of his/her own native language. This is not an exceptional situation for Northeast India, where there are several universities with linguistics departments and therefore a number of trained linguists who speak one or more Tibeto-Burman languages. For example, the linguistics faculty at Manipur University works primarily on topics in Manipuri grammar. Both the native native-speaker and the linguist native-speaker “contribute to solve the linguist’s professional problems” (Ehlich 1981:155), but the data from each group will be different. The linguist native-speaker has spent time thinking about the very issues on which the fieldworker is trying to collect information, and thus might provide only those data which support or seem important to a particular analysis. On the other hand, it is quite exciting for a fieldworker and a linguist native-speaker to set out together to discover something that has been puzzling both of them.

7.11 Third Party Interpretation

The fieldworker may find an assistant who can speak English and/or another lingua franca, as well as the target language. This assistant can then act as an interpreter, so that many more speakers can be added to the consultant pool. Since the interpreter will be present at all sessions, that person must be well respected in the community (Abbi 2001:80). When using an interpreter, it is crucial to remember how difficult translation is and how easily either the interpreter or fieldworker can misunderstand what needs to be translated.

7.12 Payment and Gifts

Much has been written on payment to consultants. See, for example, Samarin (1967b:23–27), Dimmendaal (2001), Bowern (2008:162–163), Crowley (2007:30–32), and Dywer (2006:57–60). There are three main concerns: how much to pay, when or whom to pay, and what payment means for the long term relationship between the researcher and the speaker or community.

It is fairly easy to judge how much to pay in the North American context, because many Native American communities have histories of working with linguists and anthropologists. In these instances, speakers tend to set their own rates. The fieldworker can judge if rates are reasonable by checking with other linguists who have previously worked in the area. Sometimes the requests seem inordinately high–double pay for weekend work, for example. Such requests can be accepted if feasible; if not, the fieldworker should politely point out that budget limits cannot accommodate that pay rate.

Two recurrent recommendations are (1) don't pay too little, and (2) don't pay too much. In order to not pay too little, the fieldworker could find out what speakers make at their regular job, and pay at the same or a slightly higher rate. Or the fieldworker could pay the rate which is paid for comparable services provided in the area – for example, what a school teacher or translator makes. Healey (1964:345) suggests paying as much as that person would have made had he or she been at their regular job. The hours counted for pay should include the actual hours spent with the fieldworker, plus any additional hours consumed in related travel and coordination.

On the "don't pay too much." side, one concern is that researchers not pay so much that it makes it hard for other researchers to compete with that pay rate (Dimmendaal (2001) and Genetti p.c.). Also, payment should not be so high as to embarrass or anger local employers. One aspect to consider is how the local economy is changed by bringing in large sums of money, and how some individual's lives are completely changed by becoming part of a linguist project (see Mc Laughlin and Sall 2001:209). I [Chelliah] have never had the misfortune of working with someone who boasted about how much he or she was making by working with me. I would have encountered no end of trouble if this had happened when I was in the field in Manipur, perhaps similar to the hypothetical scenario posited by Dwyer (2006:58): Suppose the fieldworker pays one speaker, and a more prominent member of the society finds out. If that prominent member is not a part of field team, there may be strong resentment against both the fieldworker and other field consultants. Speakers may also question how grant money is being spent and why some speakers receive more than others. Dwyer recommends that if "there is enough trust between you, share the project budget with the team member and explain the allocation" (58). If that is not possible, it might be necessary to increase speakers' salaries so that they are more or less equal.

One hard-to-follow piece of advice is: "Don't work with speakers who want too much pay!" A linguist might feel justified in paying more than expected if there are only a few speakers of a language left. However, it is always important to remember that the fieldworker pays speakers for their time and not for their language forms.

The recommendations listed above have always been difficult for me [Chelliah] to put in practice for my work in India, because I have rarely worked with speakers who were employed full-time. I usually set a daily rate, keeping in mind my long-term budget. I also pay for all food and travel when my consultants are working with me. It is difficult to predict how grant reviewers will react to projected pay rates in Asia: if one pays using US standards, the rate is much too high; if one pays using Indian standards, the rate seems low to reviewers unfamiliar with the Indian context. Therefore, one needs to have a reasoned approach to setting the pay rate and one needs to explain this clearly in any grant proposal. For example, I can estimate how much a university lecturer with the same educational background as my consultant might make, and then pay at that rate.

Another common suggestion is to treat language consultation as work, making clear that it is so by connecting payment with work. Bowern says, "Appropriate compensation can formalize the relationship between researcher and consultants… It encourages taking the work seriously, and viewing the work as a job that needs to be done…" (2008:163). In my [Chelliah's] experience, not everyone sees a connection between money, work, and duty to work. I am not an anthropologist, nor have I discussed this world view with my consultants, but it has become obvious to me that social obligations are far more important for the people I work with than work obligations, and this is regardless of the money involved. This is not true for all individuals, but it is true for many. Thus it should not be a source of bewilderment or frustration that a consultant would choose to work on digging a pond at a neighbor's house for no pay rather than work with a fieldworker for 2 hours for very good pay. This attitude towards work and money is related to how money is seen as a means of meeting social obligations. In many cultures, if a community member has some money, and sees a family member or friend who needs that money, that person should gladly give up that money for the greater good. Another example is given by Dobrin (2005) who describes how in Melanesia the exchange of goods is expected and is instrumental in forging meaningful relationship, allowing both parties in the exchange to be givers and takers. In terms of fieldwork then, the fieldworker should be ready to accept "gifts like strange foodstuffs…so that the bases of reciprocity are expanded" allowing for mutual dignity in the relationship (Dobrin 2008: 318). McLaughlin and Sall (2001:196–200) for further examples of consultant's attitudes towards money and work in the field.

In order to be reimbursed for payment to consultants, researchers need signed receipts. Many times the exact services along with dates and times of services rendered by the consultant must be stated. This kind of exchange – work, payment, receipt for payment – is often too formal for a particular data-gathering situation, such as when a traditional story is recorded from an elder during an informal visit. In many cultures, it would be rude to pull out a receipt book in this situation. Some options are to leave money in an envelope or inside a gift bag which also contains a small gift, or to simply leave a nice gift. Reimbursement for these expenses is tricky and perhaps impossible. The fieldworker could keep a list of such incidental expenses and try to recoup the amount from per diem payments, which are usually higher than needed

(Peter Ladefoged p.c.). What to give as a gift is also a tricky question: food is always acceptable, but what about alcohol and tobacco? See Dixon (1984) and Bowern (2008) on this point for the Australian context. Crowley (2007:32) also points out that in some cultures, "gifting" is always a reciprocated activity. Here, the fieldworker must be careful not to obligate speakers to give as much as they receive.

For speakers who are too shy to accept payment for services but who could obviously use the money, it might help to state that the money comes from a grant meant specifically for this purpose and not from private funds. Crowley adds, "To signal the official – rather than personal – nature of the transaction I [Crowley] asked [the speaker] to sign my receipt book for me." (2007:31). Crowley adds that there may also be specific customs on whether someone can accept cash payments in public. In his experience in Vanuatu, cash must not be visible when handed over.

When working with a smaller community, payment to an individual may be inappropriate. In these cases, it may be better if payment is to the community instead. George Van Driem (p.c.), for example, told us that rather than paying for fieldwork time in villages in Nepal, he provided practical gifts that could be used by families or villages; a pig or a chicken, for example, was very welcome. Dimmendaal (2001:59) suggests paying for speaker services by contributing to a child's education or purchasing household utensils. Van Driem (p.c.) also told us about work exchange: he helped build a primary school in one village, for instance. Finally, for communities that want to prepare language materials, a fieldworker might prepare children's books or word lists, or might document cultural events.

Apart from payment, speakers may expect fieldworkers to bring gifts. In the recent past, before India opened up its market to international trade, anything imported – from soap to socks and electronic gadgets – was highly prized. Today too, some everyday items can make good gifts. Certainly, the fieldworker should bring gifts for their consultants. There is a difference, however, between the fieldworker deciding what gifts to give and the fieldworker becoming an endless source of outside goods to the community. The fieldworker should decide where the limit is as to how much he or she can give in gifts. See Crowley who errs on the side of generosity (2007:32). Some speakers ask for smaller items like clothes, shoes (I [Chelliah] have several foot tracings!), and jackets. More important items that we know have been bought as gifts for consultants are bullet-proof vests, laptops, army knives, headlamps, phone cards, and books. Sometimes speakers may ask for a laptop computer. The researcher must judge which gifts are possible and which are too expensive or too difficult to carry. In some cases the requested items may be illegal: for example, most Native American Indian reservations in the U.S. are dry, and so bringing in alcohol would be illegal.

7.13 Keeping Track of Consultants

The longer the fieldworker stays in the field, the greater the number of native speakers who become involved in the project. It is easy, over the years, to forget details about consultants. For this reason, making notes on basic information for each consultant is absolutely necessary. All of this information is also needed for the metadata used in archiving. We suggest that the fieldworker maintain a separate consultant journal where distinguishing characteristics about the speakers are noted along with the following information:

- *Names*: The speaker's name, clan, caste, family name, and nickname. Here it might also be useful to assign the speaker name an abbreviation, for example, Rex Khullar would be RK. Bowern (2008:58) notes that the fieldworker should also list any pseudonyms used in published materials here. It is also useful to note how speakers identify themselves. In Manipur, for example, it is usual to give a neighborhood or village name in self-introduction.
- *Background Information*: The fieldworker should note social status, age, occupation, education, marital status, and the names of family members, if offered. (A nice early activity is to ask speakers to introduce themselves and say something about their family members. This information is helpful so that greetings and gifts can be personalized.)
- *Contact Information*: If at all possible, the fieldworker should get a phone number and address. In India, speakers will have a phone number even if they don't have an email address or mailing address. An address will be needed almost immediately after meeting the speaker in order to send a thank you note, photograph, or gift. It is polite and a good idea for the fieldworker to stay in regular contact with his/her consultants.
- *Linguistic Background*: This includes information such as: dialect, relationship to other consultants, other languages spoken, fluency in the contact language, fluency in the target language, approximate level of education, job, languages spoken in the home or by other family members, and marital status (relevant if the spouse is a native speaker of different language than the consultant).
- *Clues to Associate Names to Faces*: Keep pictures with names and addresses of consultants for future reference. I [Chelliah] met, talked with, and recorded so many speakers in my initial 9-month stay in Manipur and in subsequent visits that, 25 years later, a picture with an associated name helps me recall situations where a recording took place, etc. It is possible that no photograph, address, or name is available because the fieldworker has forgotten the camera or is restricted from taking a picture, or because no name is given at the time of data collection. A supplementary paper journal is useful for this, as the fieldworker can jot down some ways to remember this speaker: for example, noting that this consultant was a singer, or demonstrated a weaving technique might help in later recall. Similarly, when data is obtained from a speaker that the fieldworker meets for only a few minutes, some way to identify that "accidental" consultant should be found.
- *Place Where the Fieldwork is Conducted*: To help recall the circumstances under which a recording takes place, note where the fieldwork was conducted (Paul 1953:449). Healey (1964:5) says that it is sometimes not enough to get the name of the person and the village; it can also be vital to record some

geographical information, such as the nearest water sources and highest mountain. This is because village names can change. A picture might help here too because, as Healey notes, two streams which are tributaries of the same river can have the same name, as can mountains and rivers.

This information is for the fieldworker's use, and will not necessarily be made available to the public through, for example, an online archive. One should only make public the information that speakers agree can be made public.

7.14 Factors Influencing Consultant/ Fieldworker Rapport

We suggest using common sense when trying to get along with speakers in other communities. There is no guarantee that the fieldworker will "make friends for life" (Abbi 2001:80) with his or her consultants – whether or not that happens depends on the fieldworker's personality – but efforts should be made to not antagonize, shock, or hurt individual speakers or communities. If one gets the impression that the consultant does not like the fieldworker, and/or that the fieldwork or is uncomfortable for whatever reason, sometimes it is better to pay the consultant and then let him or her go even if the work has not been completed. It is better to get less for one's money than to continue in an uncomfortable situation.

The fieldworker should be honest about objectives and should clearly explain them, in order to earn and maintain the trust of the community and of individual speakers. Maddieson (2001:215) points out that phonetic fieldwork is easy to explain since speakers grasp quickly what it means to study the sounds of a language. For other topics, it is a bit more difficult, but speakers will understand if the fieldworker says that he or she wants to learn their language, and is interested in the words and in how to make sentences. See Chapter 2 on stating the goals of fieldwork.

It is important to know what previous linguistic fieldwork has taken place in the area. Do speakers feel that they have been "studied to death"? In that case, how can the fieldworker explain how his or her work contributes to the community?

The fastest way to lose friends and make enemies is to take sides in a fight that is not one's own. Are there some linguistic controversies or language ideologies that are currently being discussed? The fieldworker should be very careful not to express an opinion about issues until the full history of an issue is known. For example, controversies over the adoption of an orthographic system for Manipuri have been debated for the past 15 years. It would be irresponsible to take sides on this issue at all, and it would be doubly so to take sides without first understanding the history of the debate. On issues where there seems to be no consensus opinion, it would be best not to weigh in unless expressly asked to do so by community leaders. Even then, the fieldworker should consider the consequences for the research and for the consultants and other people associated with the project, the reaction of community officials to the fieldworker's opinion, and the legislative or pedagogical implications of voicing one's opinion.

Similarly, it is important to be informed about the historical and geopolitical factors that have shaped the society in which the fieldwork in being conducted. What is the history of race, caste, and religion in the area? In India, there can be some animosity towards the fieldworker if he or she is a Hindu working with a minority Christian community. The fieldworker may be treated with disdain if he or she is a Hindu convert to Christianity, comes from a low caste, and/or is working with high caste Hindus. There may be conflicts between speakers of the community based on similar factors. Furthermore, as pointed out by Abbi (2001:77), interaction between the fieldworker and the native speaker will also be influenced by where the studied speech community is located: urban (city, town), rural (farm, town), or traditional community (reservations, isolated and not structured). If the speech community is in an area that the fieldworker is not familiar with, there will be customs, taboos, and social rules of discourse and politeness that need to be learned (Sutton and Walsh 1979).

References

Abbi, Anvita. 2001. *A Manual of Linguistic Field Work and Indian Language Structures.* (Lincom Handbooks in Linguistics 17.) Munich: Lincom Europa.

Aikhenvald, Alexandra Y. 2007. Linguistic fieldwork: setting the scene. *Sprachtypologie und Universalienforschung.* (Focus on Linguistic Fieldwork, ed. by Alexandra Y. Aikhenvald), 60(1):3–11.

Ballmer, Thomas. 1981. A Typology of Native Speakers. In *A Festschrift for the Native Speaker*, ed. by Florian Coulmas, 51–67. The Hague: Mouton.

Barnes, J. A. 1963. Some ethical problems in modern fieldwork. *British Journal of Sociology* 14:118–134.

Benedicto, Dolores and Guillermo McLean. 2002. Fieldwork as a Participatory Research Activity: The Mayangna Linguistic Teams. Berkeley Linguistics Society 28:375–386.

Bernard, H. Russell, P. Killworth, D. Kronenfeld, and L. Sailer. 1984. The problem of informant accuracy: the validity of retrospective data. *Annual Review of Anthropology* 13:495–517.

Bhat, D. N. S. and M. S. Ningomba 1997. *Manipuri Grammar.* Munich: Lincom Europa.

Bloomfield, Leonard. 1964. Literate and Illiterate Speech. In *Language in Culture and Society*, ed. by Dell Hymes, 385–396. New York: Harper & Row.

Bouquiaux, Luc, Jacqueline M. C. Thomas, and James Roberts. 1992. *Studying and Describing Unwritten Languages.* Dallas, TX: SIL International.

Bowern, Claire. 2008. *Linguistic Fieldwork. A Practical Guide*. New York: Palgrave MacMillan.

Brumble, H. David III. 1983. Indian Sacred Materials: Kroeber, Kroeber, Waters, and Momaday. In *Smoothing the Ground. Essays on Native American Oral Literature*, ed. by Brian Swann, 283–300. Berkeley, CA, Los Angeles, CA and London: University of California Press.

Cameron, Deborah. 1998. Problems of empowerment in linguistic research. In *Le travail du chercheur sur le terrain. Questionner les pratiques, méthodes, les techniques de l'enquête*, ed. by Mahmoudian, Mortéza, and Lorenza Mondada, 23–38. Lausanne: Cahiers de l'ILSL 10, Université de Lausanne.

Cameron, Deborah, Elizabeth Frazer, Penelope Harvey, M. Ben H. Rampton, and Kay Richardson, eds. 1992. *Researching Language: Issues of Power and Method. Politics of Language*. London and New York: Routledge.

Cameron, Deborah, Elizabeth Frazer, Penelope Harvey, M. Ben H. Rampton, and Kay Richardson. 1993. Ethics, advocacy and empowerment: Issues of method in researching language. *Language and Communication* 13(2):81–94.

Carden, Guy. 1990. Polylectal Grammars, Randomly Distributed Lects and Introspective Judgements: A Replication of Legum's while-clause Studies. In *Development and Diversity, Language Variation Across Time and Space. (A Festschrift for Charles James N. Bailey.*) ed. by Jerold A. Edmondson, Crawford Feagin and Peter Mühlhausler, 205–226. Dallas, TX: Summer Institute of Linguistics and the University of Texas at Arlington.

Coulmas, Florian, ed. 1981. *A Festschrift for the Native Speaker*. The Hague: Mouton.

Crowley, Terry. 2007. *Field Linguistics. A Beginner's Guide*. Oxford: Oxford University Press.

Czaykowska-Higgins, Ewa. 2009. Research Models, Community Engagement, and Linguistic Fieldwork: Reflections on Working within Canadian Indigenous Communities. *Language Documentation and Conservation* 3(1):15–50. Online: http://nflrc.hawaii.edu/ldc/http://hdl.handle.net/10125/4423.

Davies, Alan. 1994. Native Speaker. In *The Encyclopedia of Language and Linguistics*, ed. by R.E. Asher, Volume 5, 2719–2725. New York, Seoul, Tokyo: Pergamon Press.

Davies, Alan. 2003. *The Native Speaker: Myth and Reality*. Bristol: Multilingual Matters.

Dimmendaal, Gerrit J. 2001. Places and People: fieldsites and informants. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 55–75. Cambridge: Cambridge University Press.

Dixon, R. M. W. 1984. *Searching for Aboriginal Languages. Memoirs of a Field Worker*. Chicago, IL: University of Chicago Press.

Dixon, R. M. W. 1992. Naive linguistic explanation. *Language in Society* 21:83–91.

Dixon, R. M. W. 2010. *Basic Linguistic Theory. Volume 1 Methodology*. Oxford: Oxford University Press.

Dobrin, Lise M. 2005. When our values conflict with theirs: Linguists and community empowerment in Melanesia. In *Language Documentation and Description*, vol. 3, ed. by Peter K. Austin, 42–52. London: Hans Rausing Endangered Languages Project.

Dobrin, Lise M. 2008. Discussion note: From linguistic elicitation to eliciting the linguist: Lessons in community empowerment from Melanesia. *Language* 84(2):300–324.

Dorian, Nancy C. 2001. Surprises in Sutherland: linguistic variability amidst social uniformity. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 133–151. Cambridge: Cambridge University Press.

Dwyer, Arienne. 2006. Ethics and practicalities of cooperative fieldwork and analysis. In *Essentials of Language Documentation*, ed. by Jost Gippert, Nikolaus Himmelmann and Ulrike Mosel, 31–66. Berlin: Mouton de Gruyter.

Ehlich, Konrad. 1981. The native speaker's heritage: On the philology of "dead" languages. In *A Festschrift for the Native Speaker*, Florian Coulmas, ed. 153–165. The Hague: Mouton.

England, Nora C. 1992. Doing Mayan linguistics in Guatemala. *Language* 68(1):29–35.

Evans, Nicholas. 2001. The last speaker is dead – long live the last speaker. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 250–281. Cambridge: Cambridge University Press.

Everett, Daniel L. 2001. Monolingual field research. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 166–188. Cambridge: Cambridge University Press.

Foley, William A. 1991. *The Yimas Language of New Guinea*. Stanford, CA: Stanford University Press.

Gerritsen, Rupert. 2004. Historical problems and methodological issues regarding Nhanda, an aboriginal language of Western Australia. *Anthropological Linguistics* 64(1):84–99.

Glenn, Akiemi. 2009. Five dimensions of collaboration: Toward a critical theory of coordination and interoperability in language documentation. *Language Documentation and Conservation* 3(2):149–160. Online: http://nflrc.hawaii.edu/ldc/http://hdl.handle.net/10125/4437.

Grimes, Joseph E. 1995. *Language Survey Reference Guide.* Dallas, TX: Summer Institute of linguistics.

Hale, Kenneth L. 2001. Ulwa (Southern Sumu) Places and People: the beginnings of a language research project. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 76–101. Cambridge: Cambridge University Press.

Healey, Alan. 1964. *Handling Unsophisticated Linguistic Informants*. (Linguistic Circle of Canberra Publications, Series A, Occasional Papers No. 2.) Canberra: Linguistics, Department of Anthropology and Sociology, The Australian National University.

Hoijer, Harry. 1958. Native reaction as a criterion in linguistic analysis. In *Proceedings of the Eighth International Congress of Linguists*, Eva Sivertsen, ed. Oslo: Oslo University Press.

Hopkins, Jill D., and Louanna Furbee. 1991. Indirectness in the interview. *Journal of Linguistic Anthropology* 1(1):63–77.

Kibrik, Aleksandr E. 1977. *The Methodology of Field Investigations in Linguistics. (Setting up the Problem.)* (Janua Linguarum, Series Minor, 142.) The Hague: Mouton.

Kibrik, Aleksandr E. 2007. Collective field work. In *Perspectives on Grammar Writing*: *Special Issue of Studies in Language 30:2* (Benjamins Current Topics 11.), ed. by Thomas E. Payne and David J. Weber, 25–44. Amsterdam and Philadelphia: John Benjamins.

Kroskrity, Paul V. 1993. *Language, History, and Identity. Ethnolinguistic Studies of the Arizona Tewa*. Tucson and London: University of Arizona Press.

Kroeber, A.L. 1900. *More Mohave Myths*. Berkeley, CA, Los Angeles, CA: University of California Press.

Labov, William. 1969. The logic of non-standard English. In *Georgetown Monographs on Language and Linguistics* 22, ed. by J. Alatis, 1–44. Washington D.C.: Georgetown University Press

Ladefoged, Peter. 2003. *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*. Malden, MA: Blackwell.

Larson-Hall, Jenifer. 2009. *A Guide to Doing Statistics in Second Language Research Using SPSS*. New York: Routledge.

Macaulay, Monica. 2004. Training linguistics students for the realities of fieldwork. *Anthropological Linguistics* 46(2):194–209.

Mackey, A. and Gass, S. 2005. *Second Language Research: Methodology and Design*. Mahwah, NJ: Lawrence Erlbaum.

Maddieson, Ian. 2001. Phonetic fieldwork. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 211–229. Cambridge: Cambridge University Press.

Matthewson, Lisa. 2004. On the methodology of semantic fieldwork. *International Journal of American Linguistics* 70(4):369–415.

Mc Laughlin, Fiona, and Thierno Seydou Sall. 2001. The give and take of fieldwork: noun classes and other concerns in Fatick, Senegal. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 189–210. Cambridge: Cambridge University Press.

Milroy, Lesley. 1980. *Language and Social Networks*. London: Basil Blackwell and Baltimore, MD: University Park Press.

Milroy, Lesley. 1987. *Observing and Analysing Natural Language. A Critical Account of Sociolinguistic Method*. (Language in Society 12.) Oxford: Basil Blackwell.

Mithun, Marianne. 2001. Who shapes the record: the speaker and the linguist. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 34–54. Cambridge: Cambridge University Press.

Mosel, Ulrike. 2006. Grammaticography: The art and craft of writing grammars. In *Catching language: The Standing Challenge of Grammar Writing*, ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 41–68. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Newman, Paul, and Martha Ratliff, eds. 2001a. *Linguistic Fieldwork*. Cambridge: Cambridge University Press.

Newman, Paul, and Martha Ratliff. 2001b. Introduction. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 1–14. Cambridge: Cambridge University Press.

Nida, Eugene A. 1949. *Morphology. The Descriptive Analysis of Words*. Ann Arbor, MI: The University of Michigan Press.

Nida, Eugene A. 1981. Informants or colleagues? In *A Festschrift for the Native Speaker*, ed. by Florian Coulmas, 169–174. The Hague: Mouton.

Okuda, Osami. 1995. On the Objectives of Linguistic Research on the Ainu. In *Studies in Endangered Languages, Papers from the International Symposium on Endangered Languages*, Tokyo, November 18–20, 1995, ed. by Kazuto Matsumura, *Linguistic Studies* Vol. 1, 143–148. Tokyo: Hituzi Syobo.

Paikeday, Thomas M. 1985. *The Native Speaker is Dead! An Informal Discussion of a Linguistic Myth with Noam Chomsky and Other Linguists, Philosophers, Psychologists, and Lexicographers*. Toronto and New York: Lexicography, Inc. (formerly Paikeday Publishing Inc.).

Paul, Benjamin D. 1953. Interview techniques and field relationships. In *Anthropology Today: Selections*, ed. by Sol Tax. Chicago, IL: University of Chicago Press.

Rice, Karen. 2006. Ethical issues in linguistic fieldwork: an overview. Linguistic Fieldwork Preparation. A Guide for Linguists. Online: http://www.ch2006. ass.utoronto.ca/lingfieldwork/pdf/2.pdf.

Samarin, William J. 1967a. Determining the meanings of ideophones. *Journal of West African Languages* 6(2):35–41.

Samarin, William J. 1967b. *Field Linguistics. A Guide to Linguistic Field Work*. New York: Holt, Rinehart & Winston.

Scollon, Ronald. 1979. *The Context of the Informant Narrative Performance: From Sociolinguistics to Ethnolinguistics at Fort Chipewyan, Alberta*. (National Museum of Man, Mercury Series, Canadian Ethnology Service Paper No. 52. Ottawa: National Museums of Canada.

Sutton, Peter and Michael Walsh. 1979. *Revised linguistic fieldwork manual for Australia*. Canberra: Australian Institute of Aboriginal Studies.

Sprigg, R.K. 1991. The spelling-style pronunciation of Written Tibetan, and the hazards of using citation forms in the phonological analysis of spoken Tibetan. *Linguistics of the Tibeto-Burman Area*, 14.2:93–96.

Vaux, Bert, and Justin Cooper. 1999. *Introduction to Linguistic Field Methods*. (Lincom Course books in Linguistics 1.) Munich: Lincom Europa.

Vaux, Bert, Justin Cooper, and Emily Tucker. 2007. *Linguistic Field Methods*. Eugene, OR: Wipf & Stock.

Udell, Gerald, (in collaboration with John McKenna, Sara Chapman, Francis Xavier, and Johnie D. Ragsdale, Jr.). 1972. Responses of Co-Workers to the Word Informant. In *Studies in Linguistics in Honor of Raven I. McDavid, Jr,* ed. by Lawrence M. Davis. Tuscaloosa, AL: University of Alabama Press.

Whaley, Lindsay and Fengxiang Li. 2006. The Challenge of Assessing Fluency for Documentation of Moribund Languages. Linguistic Society of America poster session.

Wong, I. F. H. 1975. Field procedures in generative grammar. *Anthropological Linguistics* 17(2):43–52.

Yamamoto, Akira Y. 1998. First Things First: on Language Preservation/Revitalization Efforts. In *Studies in Endangered Languages: Papers from the International Symposium on Endangered Languages*, ed. by Matsumura Kazuto, Tokyo, November 18–20, 1995, 213–252. (International Clearinghouse for Endangered Languages linguistic studies 1.)

Chapter 8
Planning Sessions, Note Taking, and Data Management

8.1 Introduction

This chapter reviews the steps a fieldworker should take to successfully plan, execute and document fieldwork sessions. To ensure success, each fieldwork session needs to be well thought out. We suggest the following steps:

- Create a flexible plan for each session, with clear objectives and a list of planned activities or tasks
- Find a convenient time and space for the session
- Prepare necessary equipment and other materials needed to meet the stated objectives
- Capturing session data for further analysis by appropriate note-taking, recording, data organization, and archiving practices

The primary responsibility for the success of a scheduled fieldwork session lies with the fieldworker.

8.2 Meeting Speakers, Selecting a Site

To begin with, the fieldworker will need a native speaker or group of native speakers and an appropriate place to work with the speaker(s). The uniform advice from seasoned fieldworkers, whether in a rural or urban setting, is that fieldwork should be conducted in a quiet spot where there is likely to be the least disturbance (Ladefoged 2003:21). However, the fieldworker can ask questions of consultants just about anywhere. To put a 2 hour cross-city cab ride in Delhi to good use, I [Chelliah] have pulled out a notebook and started asking questions in the taxi.

The problem, of course, is that recording cannot take place just anywhere. See Section 7.6 and Chapter 10 for suggestions on creating the optimal recording environment. Some advice based on which part of the world the research is conducted in is provided in Chapter 10. Abbi (2001) advises against working in public places like the many tea stalls that dot Indian rural and urban roads – too many "helpful"

S.L. Chelliah and W.J. de Reuse, *Handbook of Descriptive Linguistic Fieldwork*,
DOI 10.1007/978-90-481-9026-3_8,

voices may prove to be a hindrance in the initial stages of investigation. Of course, at a later stage, roadside restaurants may be the perfect place to get data based on collective discussion of forms. Eagerness to provide a "helpful repetition" as a bystander is to some extent a cultural trait. Crowley (2007:94–95, 106), while working on Vanuatu, experienced the problem of too many people repeating the same word, meaning that he could not hear anyone's repetition accurately. But Native Americans from the Southwest tend not to volunteer a repetition when they are bystanders, unless they are asked.

Dixon (2010:317) notes that aboriginal people in Australia were uncomfortable working in the "white" guest house where he was staying, and so in order to do fieldwork, he needed to go to the areas they customarily occupied.

8.3 Anatomy of a Linguistic Fieldwork Session

Wherever it occurs, the fieldworker should be prepared for a structured session. The following structure seems useful:

- A warm-up stage
- A selection of tasks
- A closing stage

A predictable internal structure to each of these stages helps the fieldworker keep track of information, such as cataloging data that he/she must routinely collect. A predictable structure also helps the native speaker judge how a session is progressing. Both fieldworker and consultant get a feeling of accomplishment as the session goes through its stages.

The steps for session preparation that we have outlined here are for the novice fieldworker; they will be modified as the fieldworker grows accustomed to particular consultants and to the pace of session activities. Bowern (2008:49) makes this nice analogy: preparing for a fieldwork session is much like preparing for a class. Novice teachers often have to write down everything they intend to say and do in class. As teachers mature in their craft, planning becomes easier and very little needs to be written down.

8.3.1 Warm-up

Each session should begin with a conversation that sets the native speaker at ease and helps bring both the fieldworker and native speaker into the artificial world created by the fieldwork session. The native speaker needs this time to block out the many distractions around him/her.

This is a good time for the fieldworker to try out any conversational phrases in the target language that he or she may have learned. One useful approach is to prepare a mini conversation beforehand to try out with the consultant, resulting in a conversational language lesson at every fieldwork session. This conversation may include performing expected gestures of hospitality such as offering the consultant a glass of water. The warm-up period is also a good time to do a final equipment check:

- Is the microphone properly placed in front of the native speaker?
- Is the recorder connected to an adequate power source?
- If a backup recorder is being used, is it ready to switch on when needed?
- Are the correct notebooks ready?
- If the consultant is being paid by the hour, has the time been noted down?
- Are headphones easily available to monitor the recorder?

Someone needs to say, "Shall we start?" But the proper time to begin work on the tasks set for the day depends a lot on the culture in which the work is being conducted and on the personality of the consultant. It may take 2 or 10 min, or it may not happen at all because the native speaker has a personal issue they want to discuss. It is important not to let personal matters hijack sessions – or, at least, discussion of personal matters should be delayed until a certain number of tasks have been completed.

The next step is to briefly review where the discussion left off at the previous session, and to discuss what progress has been made with that data in the interim. This can segue into what the fieldworker wants to accomplish during the current session. When describing goals and objectives, the fieldworker should avoid using linguistic terminology during the initial stages of investigation. It is counterproductive to go into a long (or even short) explanation loaded with linguistic terminology about what is going to be covered in the session. Some consultants might feel overwhelmed and intimidated, and might feel there is no way they can help with the project. The fieldworker must instead work on subtler ways to communicate objectives and to explain the tasks set for the native speaker. For example, it would be more effective to say (1b) than (1a).

1a. I'd like to learn more about your tense and aspect system.
1b. I'd like to learn how you talk about things that happened yesterday or things that will happen tomorrow.

8.3.2 Task Selection

Session preparation crucially includes planning what information to pursue, and how to get it from the native speaker. Part of a session will be the continuation of a previous session; for example, continuing to fill out translations for a word list or finishing transcription of a text. At least some part of each session will be based on analysis of what has already been covered. The fieldworker will review notes from the previous session, conduct preliminary analyses, and prepare an elicitation schedule based on these analyses.

Table 8.1 Relative difficulty of linguistic interview tasks

More demanding	→	Less demanding
Transcription and translation of conversations		Transcription and translation of narratives
Judgments on tone and vowel quality		Judgments on consonants
Grammaticality judgments of sentences out of context, or in contexts created by the fieldworker		Grammaticality judgments of sentences in texts, or patterned after sentences in texts
Translation of words and sentences to the target language		Translation of words and sentences to the contact language

In planning for the session, the fieldworker should intersperse more demanding and less demanding task types, as illustrated in Table 8.1. Being aware of the relative difficulty of tasks is especially important in the early stages of investigation. Mithun (2001:48) notes that fluent speakers are experts in their language but may nonetheless feel inadequate as language consultants, especially if they are not literate in the target language. For this reason, Mithun suggests that sessions should begin with interactions that bolster speaker self-confidence, and one way of doing this is to begin with a simpler task.

We arrived at the ranking of tasks as most to least demanding by considering the complexity of the question, the required knowledge of the target language and contact language for task completion, and the possibility of native speaker fatigue as the task is repeated. For example, working with a conversation is taxing because the consultant must unravel conversational overlaps and interruptions for transcription. Also, the consultant must intuit speaker intentions to properly translate the conversation. Narratives are easier in terms of transcription, because a single speaker's speech, delivered at a relatively steady rate, is to be transcribed, and a single and often predictable story line is to be translated. Speaker fatigue can set in quickly when working with tone minimal pairs, where tones begin to "sound the same" to speakers due to repetition. It is much harder for consonant distinctions to collapse in speakers' perception in this way. It takes work for speakers to set up pragmatic situations for the interpretation of clauses, and to explain or have these situations explained to them. Finally, translating from target to contact language is easier than the other way around because there is no requirement that the translation be precise; that is, the speaker can offer a paraphrase with an explanation and several alternatives of equivalent translations in the contact language, and leave it up to the fieldworker to sort out the wording of the final translation. Going in the other direction, however, the speaker can get frustrated in not being able to capture the nuance of the target language while translating from the contact language. See Chapter 12 for a more detailed ranking of elicitation tasks based on difficulty, considered from the point of view of the fieldworker as well as from the point of view of the speaker.

If a consultant becomes silent because they are having a hard time understanding the fieldworker's questions or aims, or because they cannot provide translations quickly, the fieldworker should switch to a different task. Trying out different tasks will also reveal what the native speaker's strengths are as a consultant, and what s/he feels most comfortable doing. Also, in terms of setting tasks for a session, Bowern (2008:49) recommends preparing 2 days in advance in case a task takes less time to complete than expected.

It is easy, through enthusiasm, to get a particular list of sentences translated and off the to-do list, and then to push forward with translation well past the point where the native speaker wants to continue with that task. The problem with the "must complete" approach to fieldwork is that (1) it compromises the quality of the data, because a fatigued native speaker will not be as careful as an alert one; and (2) the fieldworker runs the risk of irritating the native speaker, who might then not show up next time because fieldwork seems just too tedious and demanding. So, especially in the early stages of fieldwork, when the native speaker is still getting used to the fieldwork situation, several different task types should be used.

Finally, some tasks turn out to be lemons. They just do not work at a particular time with a particular native speaker. In this case, a fieldworker must gracefully shelve that task and continue with something else. One should always be ready to move to Plan B or even Plan C during a session. There is no need to say out loud that a particular activity is not working. One could gradually wind down one activity and start up the next, all as if it were planned. See Crowley (2007:94–98) for more practical advice on task selection.

8.3.3 What to Record in a Session

How much and what to record during a fieldwork session is proportional to how robust the speech community is. For a language that is not particularly endangered – and yes, there is fieldwork done on such languages – it is not necessary to record every single word uttered at every session. Rather, recording can be restricted to narratives, conversations, and other naturally-occurring discourse; elicited sentences; word lists and other types of data for acoustic or phonological analysis; and any discussions for which taking detailed notes is difficult, such as when a group of native speakers comment on grammatical constructions. Emmon Bach (p. c.) reflects that if one records everything during a session, the result is many unusable recordings. His preferred method is to record narratives, and then to use those for further elicitation and analysis. Crowley (2007:102) also advises that it is not necessary to record every single elicitation session.

If a language has very few remaining speakers, the "record everything" rule holds. To avoid taping too much silence, Timothy Montler (p. c.) who worked with some of the last speakers of Klallam says, "I keep my finger near the pause button and use it a lot. I pretty much know the rhythm of my speakers, so I don't miss much and don't record a lot of dead air." Keep in mind that the beginning of an utterance can be missed if one is not fast enough with the pause button. Another method suggested by Alec Coupe (p. c.) is to record everything, and then later use an audio editor to delete the non-meaningful silences.

In addition to thinking about how robust a language is, the fieldworker should also consider how long they are going to be in the field. If access to native speakers is limited, everything could be recorded. In this case, it will be necessary to be doubly conscientious about cataloging the topics discussed so that the fieldworker can later make sense of the recordings. It is helpful to say what topic is being covered directly on the recording. See also Ladefoged (2003) on being a liberal recorder. Because of the cost of memory cards,[1] space on those cards, and their limited availability in the field, it is necessary to prioritize recordings. For non-endangered languages, the fieldworker might record the most archive – appropriate materials in wav format, and other materials in mp3 format.[2]

8.3.4 How to Close a Session

The end of the fieldwork session should be a recap of what was accomplished, which could be couched in thanks, as in *"Thank you, I learned a lot today. I never realized that word* x *meant* y*, etc…"*. The fieldworker might also say what will be covered at the next session. These simple steps give the native speaker a feeling of being vested in the process. A checklist for the fieldworker for the end of a session is as follows:

- Did I make clear plans on when and where the next meeting will take place?
- Did I check on contact information, e.g. a name, address or phone number?
- If payment must be made at this time, the speaker should sign a voucher for later payment or a receipt of payment. If payment is to be in gifts, then a note should be made of what has been given or needs to be given to the consultant.
- Are notes and recordings referenced with the date and name of the native speaker?
- Did I indicate on the elicitation schedule when and where to pick up again at the next session?

Once back at headquarters, the following steps should be taken: recordings should be backed up, and notebooks should be stored so that they are easily retrievable for the next session.

---

[1] Memory cards or flash cards are media for storing digital information. Examples of devices that use memory cards include cameras, game consoles, audio and video recorders, mobile phones. Recorded digital data can be transferred directly from a memory card to a computer.

[2] Mp3 files are created by compressing the sound signal. The parts of the signal that are masked or deleted are those which a proprietary algorithm determines are not perceived by the hearer. This psychoacoustic motivation to the reduction of sound is acceptable to some fieldworkers on the grounds that, "If we can't hear it, it cannot be linguistically significant." However, fieldworkers also report that mp3 files are less useful for spectrographic analysis (Bowern 2008). Matthew Gordon (p.c.) tells us mp3 format could be problematic for some kinds of spectral analysis where precise frequency information is crucial (e.g. fricative spectra). Thus even though compression makes mp3 files smaller and thus easier to transfer and store, it is preferable to use uncompressed formats.

This marks the end of the session. This general plan holds regardless of how short or long a scheduled linguistic session is, or whether the session takes place on the road, moving from community to community, or on site. The fieldworker must have a plan, and the plan must have structure.

How long should a fieldwork session last? Many of the linguists we interviewed recommend a maximum session length of 2 hours, and this seems a reasonable length for most fieldwork situations. There are some situations in which native speakers may work for longer, more intense periods. For example, in 2007, Manipuri speakers and linguists Harimohon Thounaojam and Ch. Yashawanta Singh came to the University of North Texas to serve as consultants in a Field Methods class. In the mornings they worked together with Kevin Mullin, then a graduate student, to record tone minimal pairs and to discuss tone in Manipuri compounds. In the afternoon, they worked with me [Chelliah] on case and focus marking in Manipuri. Then, in the evenings for 4 days a week, they attended and participated in the Field Methods class. Another example is work that I [Chelliah] did with a single consultant, a trained linguist, in which we met for about 5 h a day for 1 full week. We had a very clear plan of what needed to be checked as we prepared a Manipuri text collection for publication, and this made it possible for us to work at a sustained pace. Also, there was a small window of time when we both were free to do this work. This kind of checking work does not fit into the traditional definition of fieldwork, i.e. working a few hours a day with a consultant in the community where the language is used. However, the work I did on the text collection resulted in much of the same products as traditional fieldwork: we discovered new words, discussed constructions and meanings, and did some word-list recording. Even during these long sessions, I followed the basic ideas laid out in this chapter: pre-planning, structuring the session, and interspersing difficult and easier tasks. The most unproductive use of native speaker time is doing the same task for hours at a time. See, for example, Samarin's account of trying to work on ideophones for long stretches of time (1991:168).

While it is not advisable to stretch a session for many hours because it can fatigue the consultant, a fieldworker must sometimes work long days. This is often unavoidable. I [de Reuse] remember one exhausting day when I worked for 8 hours in a row with four different speakers of Siberian Yupik Eskimo, 2 hours each. Such a day is not a typical one. Nancy Caplow (p.c.) told us that on a fieldtrip to Nepal to record word lists for acoustic analysis, that there were a few days when she recorded six speakers; three in a long session in the morning, three in a long session in the afternoon. Grueling for her and her field assistant, but not too heavy for any given speaker.

When preparing for the next day's session, the fieldworker could listen to the recordings made (using headphones to reduce background noise), and make notes on gaps in transcriptions. The fieldworker should make notes of things that need to be re-checked. If necessary, the fieldworker might listen to selections at slower speed and re-transcribe where needed.[3] A different colored pen should be used,

[3] An efficient way of doing this is to use a transcription program such as TRANSCRIBER. The waveform can be easily divided into chunks and each chunk can be played back.

to make it clear which transcriptions come from work with the native speaker and which were added later by the researcher.

When using transcription software such as TRANSCRIBER, the transcription and meaning of a construction is typed below and aligned with the speech signal. It is a good idea to do this, especially when working with a tone language or a language with a great deal of homophony or polysemy, so a native speaker can check transcriptions and glosses the following day. It is very tiring to prepare for another round of sessions after a full day of work; however, preparation time affects the quality of sessions and so should not be skipped. Minimally, preparation time should be used to back up and catalog data as soon as possible to prevent data loss. See Section 8.6 for further discussion of record-keeping.

8.4 Interviewing Techniques

The fieldwork session is often considered to be a type of interview. Like other speech events, interlocutors in an interview have expectations that inform how each speech act is interpreted. The fieldworker should take these expectations into consideration when issuing requests for information.

8.4.1 Speech Acts in the Interview

The interview is a common speech event in Western culture, where the interviewer expects to control the floor or have the power to grant control of the floor, to select the topics to be discussed, and expects the interview to proceed in predictable and orderly turns at talk. The native speaker might not be familiar with the interview speech event, and may instead expect the fieldwork session to be a conversation (Milroy 1987:41–51). Since direct questions in a conversation puts the questioner in the one-up position vis-à-vis the person expected to produce an answer, the native speaker could interpret direct questions as a challenge or threat. Answers to direct questions produced under these conditions might not be reliable. The interviewee might adopt what Milroy calls "resistance strategies" , where answers may be hard to interpret, evasive, or pared down to "yes" or "no". An all-time non-favorite is the "maybe" response for a grammaticality judgment. Milroy provides these suggestions for how a fieldworker can work around this situation:

- Use tag questions rather than yes-no questions, e.g. use "This means 'egg', right?" instead of "Does this mean 'egg'?"
- Make the purpose of the question clear by providing an introduction to the topic to be investigated. This should be phrased in everyday language.
- Use simple questions rather than questions with several parts. A complex question which asks for details and background information may elicit a partial answer. See DuBois and Horvath (1992) for further discussion of this point.

- Take on the role of a language learner.
- Work with groups rather than individuals. Group discussion of questions can free shyer speakers to contribute insights in a non-threatening environment.

8.4.2 *Expectations and Demands*

Each question asked of the native speaker reflects what the fieldworker expects the native speaker to know. If the native speaker cannot produce the required information and feels pressured to come up with an answer right away, he or she might make something up, stop talking, or get defensive. For this reason, the fieldworker's questions should not imply an expectation that the native speaker be a linguist, by using linguistic terminology, for example. Also, the fieldworker's questions should not require that the speaker provide analytic or factual answers: e.g. "How do you form questions?", or "How many consonants are there in your language?", or the very tempting "Why did you say X rather than Y?" (Crowley 2007:104). As Mithun (2001:49) states, "demanding answers to such questions, even with less technical terminology, can make speakers highly uncomfortable, and can result in responses that lead the researcher down the wrong track. It can also create an unfortunate social situation in which the researcher openly rejects opinions that have been offered, on demand, by the speaker." Over time the fieldworker develops a feel for what the consultant can and cannot do, and for how abstractly a particular consultant can think about language. There may come a time – quickly with some speakers, but never with most – that the consultant can understand questions like: "What is the plural of that?", "And the past?", or even "Can you conjugate that?" (Samarin 1967:137).

Another point has to do with speaker personality. Some fieldworkers like to create hypothetical situations and use these situations as a backdrop to elicit grammaticality judgments from the native speaker (Matthewson 2004). However, there are speakers who have difficulty with hypothetical scenarios, no matter how realistic or how simple. In these cases, the fieldworker will have to let the speaker make up the hypothetical situations him/herself, or will have to elicit actual situations where the grammatical construction under consideration can be used.

Finally, the fieldworker needs to be aware of speaker practices based on culture and context; for instance, speakers may be averse to stating the obvious (Hopkins and Furbee 1991). More about "Pragmatics Police" and "Culture Police" sorts of speaker attitudes will be provided in Section 12.2.2.3.

8.4.3 *Evaluating Native Speaker Responses*

To control the quality of data, a fieldworker must evaluate native speaker responses. While the native speaker is rarely wrong in the data they provide (Mithun 2001:44), native speaker responses may be skewed due to the following factors.

- *Wait time*: Paul (1953:445) notes that silence from the native speaker does not mean they have shut down or are thinking about other things. They may well be processing the question posed to them. Also, rules of turn-taking are not the same in all languages; we know that in some languages, such as Navajo, relatively long silences are acceptable between turns (Crown and Feldstein 1985:48f). Once adjusted to the pace of a particular consultant and language, the fieldworker might reformulate a question if an answer is not forthcoming as expected. In the case of an endangered language, the speaker may need some time to think and internally hear the language again before remembering forms.
- *Fatigue*: Fatigue may result from sessions that run too long, or from repeated requests for the same type of data. Frustration can also set in due to the inability of the fieldworker to understand the use or meaning of a grammatical construction that seems obvious to the speaker. Speaker fatigue can lead to judgment fatigue (see Chelliah 2001 and Sections 12.2.2.5 and 12.4.2.3), impatience, and irritation. In terms of responses, the fieldworker should watch for one-word answers, less thoughtful responses that come too quickly, or the same answer given several times.
- *Priming effects*: We know that speakers tend to repeat structures encountered in recent discourse. This phenomenon, known as priming, can be seen in the repetition of phonological, lexical, or syntactic constructions in natural conversation, as well as in reading and writing. Priming effects have been observed between native speakers, and in the interaction of second language learners (Branigan 2007; Kim and McDonough 2008). In eliciting language data from native speakers, it is important to guard against prompting a particular construction based on the form of the question. For example, a consultant may copy the word order of a prompt even if that is not the most natural way of saying something in the target language.
- *Dressage effects*: Another issue that the interviewer must be aware of is dressage, which refers to the fieldworker unconsciously biasing the speaker's responses to favor what he or she wants to hear. This is a danger with obliging native speakers, and is especially problematic in fuzzy grammatical areas, when grammaticality judgments are not clear-cut. To prevent this unconscious training of the consultant, the fieldworker might avoid the following types of leading questions: *"This means X, right?", "It would be great if X meant Y. You see how this would fit with paradigm Z. So, does X mean Y?",* or *"Let me just ask you about this one again, are you sure it is not an acceptable sentence?"*

There is a fine line between sharing the excitement of discovery with a consultant and revealing so much of the fieldworker's ongoing analytical struggles that s/he influences the data collected. In the early stages of fieldwork it is best not to discuss abstract goals, e.g. working out the relative clause accessibility hierarchy. If speakers do understand such statements of session goals, they may unconsciously create structure to fill the goals. Consider, for example, the following hypothetical topic introduction in (2):

(2) "In English we use an *–ed* ending on a verb to show that an action has already occurred, and I wanted to find out if the same sort of thing is true for language X. So, let's begin."

If the consultant has understood the details of this statement, he or she might unconsciously make the target language fit the English model by providing a single strategy for indicating past tense, such as a past tense suffix to match the English example. There are some things that can be done to prevent such obliging responses. Nida (1950) suggests that the fieldworker not look happy when the native speaker says something that confirms an analysis or theory. Conversely, the fieldworker should temper responses so as not to show frustration upon getting an unexpected answer, when the speaker is unintelligible, or if the speaker provides information that contradicts earlier statements. Speakers can sense this frustration and might either start giving answers that they think will be appreciated, or else simply shut down. See also Section 12.2.2.5 for additional discussion of dressage effects.

- *Influence of prescriptivism*: Often a native speaker is selected or self-selects to be a linguistic consultant because he or she has an interest in language structure. The native speaker may have attended school where he or she studied the grammar of his or her native language, of a lingua franca such as Indonesian or Swahili, or of a classical language such as Sanskrit, Arabic, or Latin. The native speaker might even be a language teacher; Fiona Mc Laughlin's Pulaar consultant, for instance, is an Arabic teacher (Mc Laughlin and Sall 2001:205). In these cases, the fieldworker must beware of the influence of prescriptive standards in the data being offered. The prescriptivism may come directly from standards set for the target language, may be based on the grammar of educational language medium, such as Hindi in many parts of India, or may be based on a classical language: e.g., "Latin has seven cases, therefore I will show that my language has seven cases and is as good as Latin." Elicited data should always be checked, whenever possible, against naturally occurring speech to guard against artificially restricted structures produced through the pernicious effect of prescriptivism. See Chelliah (2001:162–164) and Abbi (2001).
- *Influence of orthography*: The fieldworker must carefully consider comments on the sound system which are influenced by spelling. Orthography is a powerful prescriptive force. If a speaker says that the target language has two different [s]-like sounds, is this really the case, or does the orthography represent one [s] sound in two ways depending on whether the word is borrowed or native, and does this influence the speaker's perception of the sounds of the target language? Also, the orthographic system may not show allophonic variation, or may neutralize the difference between two sounds, representing them as one. Native speaker comments must always be checked against the fieldworker's close phonetic transcription. This cross-checking can, in fact, be quite helpful in revealing contrasts or variation not explicitly described by the speaker.
- *Inconsistencies*: Speakers will give inconsistent judgments, or provide variants of constructions that may support contradictory analyses. The fieldworker must determine why this apparent inconsistency exists. Whalen (1981:264–275) discusses the following reasons for individual inconsistencies, which he clearly

distinguishes from variation caused by dialect differences. One reason for speaker inconsistency is that the constructions under question were elicited using faulty elicitation techniques. For example, if elicitation was based on asking for grammaticality judgments, the speaker may on one occasion have accepted a sentence as grammatical because of a plausible context for that construction in the immediate field session, but on another occasion might reject this construction because a context is not available. Another reason is that some examples are hard for speakers to commit to because their status as acceptable sentences is so marginal and the constructions are so rarely used by the speaker. A sentence may be hard to parse – e.g., a sentence with quantifier and negation such as *All the boys didn't leave* – and therefore may have varying interpretations (Whalen 1981:270). So, some speakers might interpret this sentence to mean 'All the boys didn't leave (all stayed)', while others might interpret it to mean 'Not all the boys left', and a third group might accept both interpretations. Other reasons for variation are the well-known factors of register shift and style shift. The possibility of inconsistent responses requires "getting a number of speaker judgments about any but the most trivially determinate sentences …" (Whalen 1981:275).

- *Absolutes*: Speakers may make statements qualified with "never" or "always" or "only." Again, the fieldworker must look for confirmation in naturally occurring speech. Following a prescriptive rule, most Manipuri speakers will state that in Manipuri the subject always occurs with a case marker *-nə*. Further elicitation and study of texts to show that this is not an accurate picture of subject marking (Chelliah 2009).
- *Group dynamics*: Anthropologists and language teachers have noticed that responses from native speakers in a group can be quite different from those provided by individuals. Speakers may defer to the judgment of a socially superior speaker but give a different answer when that speaker is not present. See also Section 7.6.3 regarding group dynamics. Thus while group discussion can yield interesting material (Ladefoged 2003), the fieldworker cannot be sure of the data without checking with speakers outside of that situation, and with naturally occurring data.

8.4.4 Reacting to Native Speaker Responses

We have already talked about monitoring fieldworker reactions to guard against dressage effects. It is equally important not to criticize the information provided by the speaker. The fieldworker should avoid "teaching" the native speaker about their language unless explicitly asked to do so. For example, a speaker might insist on a word being transcribed in a particular way. Even if the fieldworker hears things differently, they should not correct the speaker. Rather, a note should be made of the discrepancy in an unobtrusive fashion, and the data should then be checked with another speaker (Healey 1964). The difference between a fieldworker's perception and the speaker's transcription choice usually reflects something interesting; for example, the difference between fast and slow or formal speech.

Consultants, even when regarded by the fieldworker as linguistically unsophisticated, must be acknowledged, recognized, and treated as equals. Nida (1950) and Healey (1964) discuss how easy it is to make the fieldwork session a power play, a way of showing the native speaker how bright the fieldworker is. To avoid this, the fieldworker should:

- Accept corrections or analyses suggested by the native speaker, and show interest in group discussions on language use and structure
- Never put the native speaker on the spot by asking him or her to defend an opinion or by bringing up a past opinion that contradicts something elicited more recently
- Never pit one native speaker against another
- Frame reactions or responses to native speaker contributions so as not to embarrass or offend the speaker
- Not require too much independent work from the speaker without knowing that they can deliver
- Create an atmosphere of mutual respect, as in acknowledging the speakers' contributions to the project with comments like: *"This will really help us…"*

8.4.5 Maintaining Native Speaker Interest

If the planned fieldworker session is essentially a Question and Answer session, the native speaker is likely to get bored and disinterested. While it is important for the fieldworker to keep control of the session and complete specific work (Ladefoged 2003), it is equally important to give the native speaker a chance to provide unprompted input. Here are a few suggestions on how to build speaker-led interactions into the fieldworker session:

- Allow time for conversation using the strategy of "associative interrogation" (Kibrik 1977:58). Abbi (2001) provides the following use of this strategy as an example: eliciting the word *chappati* 'Indian unleavened bread' leads to a conversation about how the bread dough is prepared and cooked. The discussion may provide previously unknown constructions and words; help to fill paradigm gaps; and/or reinforce previous analyses. Associative interrogation allows the fieldworker to relieve the monotony of elicitation and collect information that he or she may not have thought of eliciting or may not have come across in texts.
- Speakers should be encouraged to supplement answers to direct questions with further examples and illustrations. Abbi (2001) suggests providing speakers with paper and pencil or slate and chalk. She notes that in rural India, women often draw illustrations of words being discussed on their earthen kitchen floors.

- Always plan several different types of tasks.
- With some consultants, the use of amusing scenarios or quirky or funny vocabulary items might work to maintain speaker interest, as further discussed in Vaux and Cooper (1999:18) and in Section 12.2.2.9.

8.4.6 Maintaining Roles During the Session

Master teachers retain control of their classroom to ensure productive use of time, but at the same time to encourage a collaborative atmosphere. A fieldworker must do the same at each session. The problem is that the longer a fieldworker stays in the field, the harder it becomes to assert and maintain strict fieldworker–consultant roles during sessions. Some possible scenarios of role change are as follows.

Sexual interest develops between the fieldworker and speaker. In most parts of India – as in other conservative societies – there would almost certainly be negative consequences for a fieldwork project if such a relationship were to develop. After having invested time, money, and effort in getting to the field site, it seems a shame to jeopardize one's relationship with the community because of a romantic or sexual liaison. If, however, the fieldworker should find himself or herself in such a situation, it is best to think beforehand how others in the community will react, keeping in mind that in small communities, just about everything one does is known to everyone. See Vaux and Cooper (1999:16) for some suggestions on this point.

Another possibility is that the fieldworker becomes a life counselor to the native speaker. It is problematic if discussion of life problems overwhelms each session. A monolingual approach at this stage might be useful, since one could at least use the time to learn more about interactive speech, such as the correct responses in what Tannen (2001) calls "troubles talk".

What should the fieldworker do if a speaker wants to "teach" the language and comes to elicitation sessions armed with grammar books, dictionaries, and the like? There are two different opinions about this. One opinion is that the teacher role for the native speaker consultant should not be encouraged (Vaux and Cooper 1999:51). This may be the correct opinion in the context of a field methods class, but it seems to us that a very useful skill would be to allow consultants to be teachers while at the same time guiding what is being taught. The fieldworker should make copies of all materials that the native speaker brings to sessions; welcome all they have to offer, and show interest in it. Prescriptive grammars contain data that can be used as a springboard for further discussion. Discussion about formal grammar instruction can also reveal prescriptive influences on the native speaker. At some point the native speaker will run out of materials to share. At this point, the fieldworker will be in the clear to continue work as planned, having learned a lot about the language and about existing materials on the language in the meantime. See Hopkins and Furbee (1991) for further discussion of this point.

8.5 Note-Taking

A field session should always involve note-taking; fieldworkers must both digitally record and transcribe answers to elicitation prompts. Digital recordings should not be used as the only method of documenting a data-gathering session. There are many reasons for this: for example, when transcribing with the assistance of the native speaker, doubts about transcription can be addressed immediately; during fieldworker–speaker interactions many ancillary but interesting linguistic and cultural facts are unearthed; and recordings can be erased or lost, and if this happens, the fieldworker will be left with no record of the session.

Notes should also be taken when transcribing and translating previously recorded speech. In the case of writing down answers to elicited materials, there will typically be one written form for several digitally recorded repetitions of the word or clause under consideration. The fieldworker should not try to transcribe every repetition as this will tire out the speaker and slow down the elicitation process.

How a fieldworker takes notes during a fieldwork session greatly influences how well the data can be accessed for later analysis.

All the comments made by the native speaker, regardless of whether the language is endangered or not, should be written down. The best consultants offer unsolicited information and comments on extra-linguistic factors such as orthography, language history, and language politics or policies. These comments are always useful even though they may not seem immediately so.

Traditionally, field notes are recorded in notebooks. We share here suggestions from fieldworkers on notebook quality and quantity and related matters.

Use acid-free paper, since other paper browns and disintegrates in a matter of years. (Since bound notebooks with acid-free paper are costly, this expense should be factored into grant budgets.) In some parts of the world, acid-free paper is not easy to find. If it turns out that a low quality notebook must be used, handwritten field notes should be digitally captured as soon as possible. This can be done with a digital camera, if a scanner is not available. If digital capturing is not possible in the near future, sending a photocopy of the notebook to another location is always a good idea.

Each notebook should be clearly numbered and should have a designated topic so information can be easily accessed after fieldwork. I [Chelliah] use one set of notebooks for translation and discussion of narratives, and another set for elicitation on specific topics such as tone, question formation, and verb/affix ordering. If this plan is followed, it is useful to have notebooks of the same size and quality but with different colored covers. As many notebooks as possible should be taken to the field; more is better because there should not be a lot written on each page. Also, a few extra notebooks should be reserved for use by consultants.

Sometimes notebook size matters. We recommend rather large sizes for the collection of texts, but small sizes for lexical note taking, or for unobtrusively writing down overheard expressions.

Emmon Bach (p.c.) recommends always using ruled notebooks; Bowern (2008:49) recommends bound notebooks that can be zipped up in plastic bags.

This, or plastic covers, or leather binding, will help in damp climates. Bowern adds that a strong binding will also help if the notebook needs to be balanced on the knee when writing. Of course, bound notebooks are bulky and heavy and will make carry-on luggage that much heavier.

The ink used should not easily run or smudge, so pens should be selected with this in mind. The ink should be black, since other colors do not photocopy as well. The paper should be thick enough so that pen does not show through on the other side. Pencil fades and should not be used when taking field notes. (In any case, it is not recommended that transcriptions be erased.) Often, first guesses at transcription are correct, a fact that may only be discovered after the fieldworker meanders in the transcription wilderness for a while. To preserve these early guesses, Bowern (2008:50) suggests that if a mistake in transcription is made, it should be crossed out with a single stroke so that the "incorrect" transcription can still be accessed. Furthermore, early variations in transcription may indicate allophonic variation; an allophone may have been heard at an early session but not in later elicitation, so the first transcription may be the only record of it. Also, having access to early mistakes will help keep track of how transcriptions and analyses have changed through the course of fieldwork (Abbi 2001).

It is very important to put some thought into how fieldwork notebooks are going to be organized. Two main lines of advice from seasoned fieldworkers are (1) the less put on a page the better, and (2) space on each page should be left for cataloging and referencing information to help with data retrieval. Here are some details.

It is useful to make the first page of the notebook a Table of Contents. Here the main topics covered or texts translated along with the page numbers where the topics are discussed can be recorded. Also at the beginning of the notebook, a page should be set aside for an explanation of all the unusual phonetic or orthographic symbols or short-cut notations used. For example, in the course of elicitation, if the appropriate IPA character for a sound cannot be recalled, a substitute symbol might be used as a place holder. This character can immediately be recorded on the dedicated Symbol Page, with a note about what it stands for.

Each fieldworker organizes notebook pages differently, but here are some common guiding principles:

- Dedicate the odd-numbered pages to writing down elicited constructions or transcription and translation of texts.
- Save the even-numbered pages for incidental or related information.
- When transcribing texts, skip a couple of lines so that blank lines remain for word and free translations.
- Number the transcribed text lines.
- Transcribe one clause, one phrase, or one intonation unit per line, even though the line will not be filled all the way across the page. This may seem like a waste of space on the page, but this division will help with translating texts at a later date, as well as forcing early hypotheses on constituent structure.

The only problem with using so much notebook space for texts is that, when rapidly transcribing a text, the fieldworker might not easily see the flow of the text in long sentences or paragraphs without having to leaf back and forth quite a bit. The only way to alleviate this problem is not to use notebooks that are too small in size, and to retype the text into one's computer as soon as possible.

The same page layout can be followed for eliciting sentences from an elicitation schedule: transcribe on the odd numbered page and put incidental comments and cross referencing information on the even-numbered pages, number all elicitations, and leave space for translations if they are not immediately provided.

Next we address the cataloging and cross-referencing of information that should recorded on each notebook page. Each page holds a wealth of information. Pages should be organized in a predictable way so that information is easy to retrieve. Vaux and Cooper (1999:98) suggest preparing notebooks for a session by designating spaces for writing down anticipated information. For example, when working on nominal morphology, the fieldworker expects to elicit nouns in different cases, persons, and genders. In this instance, the notebook page might look like a spreadsheet where these forms are to be filled in.

When working on translations of conversations or narratives, it is impossible to predict what grammar points are going to come up. To keep track of constructions of interest, an area of each page can have a designated spot where the topics related to the data on that page are listed. For example, if the page contains a discussion of a negative passive sentence in the past tense, PASSIVE, PAST and NEGATIVE could be noted on the upper right hand corner of the even-numbered "extra notes" page. A space can be designated for jotting down how many seconds into the recording each line of transcribed text represents.

In terms of numbering, each entry in the notebook should have a unique identifier, such as a notebook number, notebook page number, and example number. If it later becomes necessary to re-check the meaning of a word or look at the accompanying notes, the identifying information can be used for quick information retrieval. Since we are all human, we make mistakes when typing up our notes; therefore it is useful to be able to check the original notes without having to leaf through all our notebooks.

Finally, we strongly advise that fieldworkers write down how data was acquired. Kibrik (1977:60) notes the following possibilities: translation, offered by the speaker, discussion of text, or transformation of text sentence. This becomes an important part of the record, as it allows others to evaluate the data in terms of which part of a speaker's competence it accesses. The following abbreviations have proven useful in indicating data source:

- *el.* "elicited": Response is a direct result of the fieldworker asking for the information.
- *ov.* "overheard": Language data gathered outside of a formal field elicitation session.
- *rej.* "rejected": Form volunteered by the fieldworker but rejected by the speaker.
- *unk.* "unknown": Form obtained from other sources, but not recognized by this speaker.
- *vol.* "volunteered": Information not directly asked for, but produced out of the blue in the fieldwork setting.

In addition to using notebooks for field notes, it is common to use prepared schedules that look like worksheets or questionnaires. For example, for sentences that need to be translated, the sheets might have a sentence in the contact language and then a space for translation to the target language; for word lists, the sheets might have a column of printed words in the contact language followed by a blank column for translation into the target language, and another column for sample sentences using the word.

There are obviously advantages to using worksheet-type schedules: the session is planned out in advance, and the fieldworker does not have to spend time writing down the sentences to be translated during the session.

There are also problems with the worksheet: loose paper can be lost; it is not easily placed in a sequence with other notes because it must be stored separately; and most importantly, there never seems to be adequate space for translation and commentary on each item of the questionnaire itself. For these reasons, we recommend that, when using elicitation schedules, actual notes are still recorded in a field notebook, with cross-reference to the numbered items on the elicitation schedule. If loose sheets are used for elicitation, it will be necessary to take along folders, a hole-punch, and organizing dividers. If the elicitation is available electronically, the digital file should be referenced in the written notes.

At times it is useful to take notes directly on a computer. For example, a word list could be prepared, with words on one side of the page and columns on the other with an area for English glosses, transcription, sample sentences, and other notes. A great advantage of this is that the fieldworker can then use the field notes to build or enlarge an existing lexical database. Using a computer for note-taking is not a good idea early on in fieldwork, because it may not be obvious which phonetic characters need to be keyed in. For quick typing, a phonetic keyboard may be needed. Typing in data may restrict the phonetic distinctions a fieldworker is willing to make if only a subset of phonetic characters are easily available on the computer. A Unicode font compatible with other software to be used in the fieldwork project should be used.

Not everyone is good at typing field notes directly into a computer. I [de Reuse] tried it once, but since I am a very slow typist, it took too much time, and my frustration at typing so slowly made my consultant nervous. So I keep to handwriting phonetic transcriptions fast and furiously during fieldwork sessions, and then I leisurely re-type my notes into the computer after the fieldwork session.

It is a good idea to begin using a standard set of abbreviations for morphological glossing from the outset. This will make it easier to enter data into a database. Morpheme glosses must be consistent in order to make database searches possible. Using a standard set of morpheme glosses also helps in cross-linguistic comparison.

It is also a good idea to adopt early on some of the other conventions to be used when entering data into a database. For example, when there is a many-to-one correspondence between a morpheme and its gloss, the accepted practice is as in example (3) from Lamkang, a Tibeto-Burman language of Manipur, India. The gloss 1.AGR for '1st person agreement' has two parts to it, which are fused with the period to match with the target language form.

(3) *kíʔip*
kV-ip
1.AGR-sleep
'I sleep'

The fieldworker should refer to the Leipzig Glossing rules[4] for useful glossing rules and a list of standard abbreviations.

More practical advice regarding notebooks can be found in Sutton and Walsh (1979:8–9) and Bowern (2008:47–50).

8.6 Record-Keeping

Recording, transcribing, questioning, and analyzing – all of these activities are aided by being able to quickly access relevant information in field materials. Every book on field methods includes sections on data cataloging and management, and we suspect this is because every fieldworker has had that "oops!" moment when they wish they had kept better records. Here are some of the things that can go wrong:

- The fieldworker forgets to label an audio file, and 5 years after a field trip has completely forgotten who or what is on the tape or audio file.
- An audio file is mislabeled and cannot later be matched to the right transcription.
- A digital sound file A is named with a previously used file name B, so that the existing sound file B is overwritten.
- Paradigms are spread over several notebooks, but because the content of each notebook has not been noted down, the fieldworker must go through each notebook looking for information on a particular paradigm.
- Some notes are on the computer, some in notebooks, some on loose sheets of paper, others in a diary. Years later the fieldworker discovers that time has been spent re-eliciting information that was already elicited but hidden in a forgotten manila folder.
- The fieldworker is required to archive field materials, but because the data has not been dated, cross-referenced, or accompanied by the necessary metadata, an inordinate amount of time must be spent in recreating this information from field notes.

To avoid these and other possible issues, fieldworkers are strongly encouraged to think about how they are going to organize and catalog their primary data before they head for the field. The fieldworker should aim to create an "apparatus" which will allow for efficient access of materials during data collection and data analysis, and for purposes of data archiving (Himmelmann 2006:21).

[4] http://www.eva.mpg.de/lingua/resources/glossing-rules.php

8.6.1 Keeping Track of Audio Files and Associated Metadata

After each fieldwork session or at the end of a day of fieldwork, some time should be spent organizing and cataloging audio and video files collected that day. This should be done in four steps:

Step 1: Digital capture and file naming
First, attention should be given to the digital tapes or digital recordings that have been made during the fieldwork session. The fieldworker should make backups of the tapes or files onto a computer, and from there onto CDs, DVDs, or a dedicated external hard drive. This is an easy task, but it must be done carefully, and only when the fieldworker is clear-headed, because it is very easy to delete a file permanently. Each time a digital recording is made, a file with a new number is created. The recorder might name these files like this: TRK1.01.WAV; TRK1.02.WAV, TRK1.03.WAV, and so on. These files are saved in a folder which will also have a nondescript name such as PRJ for project. (The fieldworker should be completely familiar with the recorder and the software that it comes with before the first recordings are made.) The easiest procedure is to transfer the folder from the recorder to a computer and give the folder name a date extension such as PRJ1_2009_01_18, for file of project 1, recorded on January 18, 2009. Proper naming procedures are very important; with audio and DAT tapes, the simple procedure of popping the tabs on a tape prevents recording over the tape, but digital recordings are easy to delete accidentally.

Step 2: File backup
The folder should be backed up, i.e., copied to an external hard drive or burned on to a CD or DVD. This is a very important step because it ensures at least one safe copy of the primary recordings. Archive copies should be stored in a different place from the original, so that if one batch gets damaged or stolen, there is a better chance of having a useable copy. Needless to say, each file and each piece of recording media (i.e., minidisk, memory card, cassette tape, etc.) should be carefully labeled. Austin (2006:89) cites what he calls, "a widely agreed upon mantra, LOCKSS, 'lots of copies keep stuff safe". Nothing should be changed on this folder – either the copy on the computer, the one copied to external media, or the archive copy. See Bowern (2008:53) for other choices on how to number or name files.

Step 3: Renaming working copies
After safely storing an archive copy, attention should be given to the audio files on the computer. The original folder, the one with the digital capture from the recorder, can be copied; the copy can be renamed PRJ.Wk.1_2009_01_18 (Wk for "working"). Within that folder, it might be useful to edit the names each of the WAV files so that they can be identified by speaker or content. So, for example, TRK1.02.WAV might be RobertPearStory.wav. Renaming the files in this way makes it easier to retrieve these files for use with transcription or speech analysis software.

STEP 4: RECORDING METADATA

In the notebook(s) where recorded material is discussed, the fieldworker should write down the digital audio or video file name. She or he should keep a master list of where each file is located, noting what is on each file, and in which notebook the contents of the file are discussed. The master list should be updated soon after a session. Here is a list of information that could be included. Other sample lists for metadata are in Samarin (1967), Ladefoged (2003), Bowern (2008), and Himmelmann (2006).

- Unique identification number
- Date, time, and place of recording
- Name of the person who did the recording
- Topic of recording
- Person(s) being recorded
- Specifics of equipment used: microphone, recorder, video recorder
- Indications of whether permissions are oral (recorded on the sound file for illiterate participants) or written (signed IRB agreement forms)
- Cross-reference to transcriptions
- Sampling rate and format of recording (wav, mp3, for example)

Innes (2010:2–4) suggests that the following types of questions on data (e.g. a story) would help future linguists know how to appropriately use a text:

- What time of the year can the story can be related and studied (In some cultures certain stories can be related only during some parts of the year or day)?
- Was the story related over the course of several days or in 1 day?
- What is the suitable audience for a story? (e.g. in some cultures, men should not hear childbirth stories)
- How should storytelling interactions be begun or ended?
- What ideologies and practices determined the telling? (For example, is the story known to the entire community, considered scary, playful, representative of a stigmatized or prestige vvariety?)

A useful common practice, as discussed in Ladefoged (2003), is to voice-record metadata information at the beginning of each recording. Also, at the end of the recording, the fieldworker can summarize what has happened, especially if there were any changes from the original plan, e.g. the speaker told two stories rather than three, or the speaker gave more words than were on the list.

When doing fieldwork on Lamkang, I [Chelliah] transferred files from my digital recorder to my computer at the end of the session before packing up my equipment. Then, once I had some down-time, I labeled the digital folders according to the date they were recorded, and copied them to a safe location. I have found it very useful to keep a set of notes in a fieldwork journal where I record exactly who I worked with on a given day, where the work was conducted, if I took pictures of the consultants, who accompanied me to the fieldwork site, and what the content of recording was. Also, this journal includes notes on conversations I had with my consultants that day: if someone tells me about an uncle who knows traditional stories, I write that down;

if someone tells me the name of a speaker living in the United States, I make a note of that as well. This journal can be a useful tool for reflection, and keeping such a journal has all the benefits associated with journaling (Cooper 1991).

8.6.2 Keeping Track of Field Notes and Associated Metadata

At the beginning of each session, the fieldworker will also need to record information that will help cross-reference field notes with associated audio recordings. This information includes details about the session and speaker. There are several discussions in the literature on what information to record about the data collected during a session (session metadata or item metadata) and information collected about the participants (speaker metadata). Below we provide a sample list of item metadata, and refer the reader to Section 7.13 for a list of speaker metadata information. See also Bowern (2008:57) and Good (2002) and the links found there, and the DoBeS Archive (2006)[5] for more on the reasoning behind recording metadata. Austin (2006:93) categorizes metadata into these five types:

- Cataloging: Useful for locating the data and identifying who collected it and who provided it. This would include: consultants and others in attendance at the session; date of collection; target and contact languages used; where the data was collected and other details of the setting; reference to an associated sound file using the unique identification number; title of text(s) collected.
- Descriptive: The genre of the data (e.g., narrative, speech, or conversation, for example); what topics and subtopics are covered in this session (e.g., Phonology: assimilation; Syntax: case marking).
- Structural: How files are organized.
- Technical: The software requirements for accessing files.
- Administrative: Access restrictions, intellectual property rights, general archive identifying information.

Some of the information here is duplicated from the information on the actual sound files (Section 8.6.1). This redundancy helps in correctly linking the sound files with the corresponding field notes. For most researchers, this information can be managed in an EXCEL spreadsheet, but one should keep in mind that EXCEL is proprietary software and will eventually become obsolete. Therefore, and especially if there are several consultants involved, it is better to use programs such as IMDI (the ISLE Metadata Initiative) Editor which was developed at the Max Planck Institute for Psycholinguistics, Nijmegen (http://www.mpi.nl/IMDI/tools), or the Open Language Archive Community (OLAC) for metadata recording which uses an XML format (http://www.language-archives.org/OLAC/olacms.html).

[5] http://www.mpi.nl/DOBES/

In the field and en route back to one's home institution, every fieldworker has that little nagging fear that the materials he or she has collected may not make it back safely. One absolute rule is to not put in one's checked luggage anything that is not replaceable. When going to the field, the expensive tape recorder and other equipment should be carried on the plane. On the way back, however, the tape recorder might go in checked luggage, while items that cannot be replaced, such as field notes (perhaps 10–15 bound notebooks), other printed material, photographs, flash drives, hard drives, and laptop computer should be carried on the plane and not leave the fieldworker's side.

Original field notes should never be mailed. We know of someone who mailed the only copy of her notes back to the United States from South America never to see them again! Remember that even if the mail system is dependable, one can never tell if something is going to get lost. Email, however, is a good option for keeping copies of files safe. Kristine Hildebrandt (p. c.) tells us that one way she felt safe about her recordings was to email the most important files to herself on a regular basis. If possible, photocopies or scans of field notes should be made. Nancy Caplow (p. c.) copied her field notes in the field using a digital camera. Michael Krauss (p. c.) once put copies of his notes in a safe-deposit box in a bank, and as far as he knows, they are still there!

8.7 Transfer of Data to Database

The most effective use of the information in a researcher's notes will be to bring together material on related topics in a structured and searchable database. It is true that this can be done without a computer, and previous generations used cards in shoeboxes to sort and resort data. But there are powerful tools available today to help with just this type of data organization, and with the correct preparation it should be possible to use these tools for faster and more effective organization.

A first step is to process data so that it can be successfully transcribed and translated. At the time of this writing, a common speech analysis program used by linguists is PRAAT (Boersma and Weenink 2010). TRANSCRIBER, originally developed for the transcription of broadcast news recordings, is useful for transcription of conversations. Either TRANSCRIBER or PRAAT can be used for data transcription. TRANSCRIBER, ELAN, and EXMARaLDA are useful for time-aligning speech with annotations. ELAN also allows for aligning video with audio and text, a feature that is absolutely necessary with work on sign language.

A popular piece of software – often used in conjunction with TRANSCRIBER and ELAN – is TOOLBOX, which is useful in creating lexical databases and annotating interlinear texts. The next generation of TOOLBOX, the Fieldworks Language Explorer (FLEx), is now in limited use. FLEx is like TOOLBOX in that it allows for annotating lexica, which can be used to create a dictionary and to create interlinear translations of texts. Butler and van Volkinburg (2007) discuss other features of these software packages.

To key in Unicode characters, an existing keyboard can be used, or a custom keyboard can be created using a keyboard utility such as Tavultesoft Keyman.[6]

A note of warning regarding the software discussed in the previous paragraphs: setting up software for use in a project will take more than a few hours to figure out; time needs to be designated for trial and error and for pre-planning how data is going to be entered. This should be done as early in the project as possible so that large scale changes will not have to be made midstream. Transferring files from one program to another, re-entering data using different fonts, re-doing the way data is glossed or structured, and adding fields for annotation take time and can slow down analysis considerably.

Using transcription and database programs provides fieldworkers with another housekeeping challenge: data can end up in so many different places that a systematic plan for numbering and naming files must be followed. Take, for example, the types of information that can be created for a single narrative:

- A sound file stored on the computer, perhaps in more than one format.
- A phonetic transcription found in a notebook or computer file. Let's call this the "Transcription file". It might be generated with software such TRANSCRIBER or simply typed into a word processor.
- An orthographic transcription file provided on paper and then scanned in, typed in, or provided in an electronic format.
- A file with free and word-for-word English translations. These could be found in a notebook and/or a digital file. The "Translation file" may be the same as the "Transcription file" if it is generated by software such as TRANSCRIBER or ELAN.
- A file with free and word-for-word translations in the lingua franca.
- There may also be a morpheme-by-morpheme analysis that is gradually being filled in. Most likely, this is part of a database file, generated from software like TOOLBOX or FLEx.

The individual programs will also generate their own files associated with each annotation project that is started. There must be a plan regarding how this information is to be organized. A sample file structure for a language documentation project is outlined below:

1. **WORKING FILES**
    a. PRIMARY DATA
        i. *Audio and Video Files*
            1. By date
                (1) By topic
                    i. By track

[6] http://www.tavultesoft.com/

ii. *Notes Files*
  1. By topic (e.g. relative clauses)
    (1) By session (e.g. January 15, 2009)
      i. Organization based on preliminary analysis (e.g. relativized direct object, relativized subject, etc.)

b. Organizational Data
  i. *Metadata*
    1. Consultants
    2. Project
    3. Access
    4. Item

  ii. *Administrative files*
    1. Receipts
    2. Expense records
    3. Work Schedules

c. Speech Analysis Files
  i. *Data analyzed with PRAAT*
  ii. *Data analyzed with TRANSCRIBER*
  iii. *Data analyzed with FLEx*
  iv. *Data analyzed with Elan*

d. Software
  i. *PRAAT*
  ii. *TRANSCRIBER*
  iii. *FLEx*
  iv. *Elan*

2. **PRESENTATION FILES FOR PAPER OR WEB PUBLICATION**
  a. .pdf files
  b. .html files
  c. .doc files

3. **ARCHIVING FILES**
  a. .xml files
  b. Metadata files
  c. Scanned field notes
  d. Original audio and video files

As shown above, there are three higher level directories: one with the working files; the second with presentation files for paper and web publication; and the third with files for archiving. Another higher level directory may be created for community outreach materials such as multimedia products, dictionaries, and community training materials on orthography and documentation, as discussed in Nathan (2006). For another example of file organization see Drude (2003).

8.8 Transfer from Database to Archive

Austin (2006:100) explains digital archiving as "the preparation of the recorded/captured data, metadata, and processed analysis so that the information it contains is maximally informative and explicitly expressed, encoded for long-term accessibility and safely stored with a reputable organization that can guarantee long-term curation."

Why bother to archive if the primary goal of descriptive fieldwork is data collection and analysis? If data is stored in a suitcase in the researcher's attic or office, that data is in danger of being lost. It is certainly not being put to any use. Although a data archive is not a backup system, archiving data is a safe way to house recordings and analysis for the long term. Additionally, in a digital archive, language data can be made accessible to native speakers and interested linguists with the added advantage that while all appropriate materials are housed together, not all need be universally accessible. As Himmelmann (2006:6) points out, an archive also provides data against which other linguists can check the veracity of a researcher's claims.

When doing fieldwork, the fieldworker should think about the information that needs to be put in the archive so that the appropriate metadata can be collected. For archiving purposes, metadata of the following types should be noted:

- Type of material: (e.g. picture, text, interlinear translation, or elicited sentence)
- Genre (e.g., narrative, speech, or conversation)
- Summary
- Recording specifics (e.g. recording equipment used and sampling rate)
- Format (e.g., wav, Excel, Word, Transcriber, or Elan file)
- ID/Filename (e.g. name and location of file names, name and location of transcripts, name and location of field notes)
- Rights management
- Speaker information (e.g. name, dialect, age, picture)
- Length/pages (byte size or page number)

The Dena'ina Qenaga Digital Archive[7] is an example of an archive for a single language, Dena'ina, an Athabascan language of Alaska. It contains sound files and documents. For each sound file and document the following information is provided: Identifier, title, date, and participant.

Another type of archive might include all the documents from a project, including: scanned field notes, sound files, previous data, scanned published material that is not easily available (if there are no copyright restrictions), analyzed data, and all other supporting documents such as photographs. Archives containing information on several languages with multiple contributors are PARADISEC,[8] described in Thieberger (2005) or AILLA, the Archive of Indigenous Languages of Latin America.[9]

[7] http://qenaga.org/index.cfm

[8] http://www.paradisec.org.au/home.html

[9] http://www.ailla.utexas.org/site/la_langs.html

Information needed for a complex archive such as AILLA includes: depositor information, specifying the role of the depositor (e.g. interlocutor or compiler); agreement forms showing that the depositor has permission to share the deposited data; collection data, which is similar to the session data described in Section 8.7; and access conditions, such as who can have access to the texts and which texts should be freely available on the web. Such a website may also include what Himmelmann (2006:13) calls “general access resources” which would include: an overview of the documentation and description project; background information on the language and speech communities; notes on practical orthographies and transcription conventions; a grammatical sketch; and references to related resources. Older documents to be used in philological study can also be included in the archive, as in the case of AILLA.

Much of the material to be archived may be handwritten, in which case it must be rendered digitally in a format which is machine readable and can be marked up. An example of digitization of handwritten field notes is Henderson (2008)

Data should be accessible to a wide array of people many years down the road (Woodbury 2003). Therefore, information must be stored in a format that can be read without access to proprietary software. Currently, it is recommended that text materials be archived in XML format, in which data structures are overtly stated rather than encoded in a proprietary format which will no longer be readable as soon as the required software becomes unavailable (Austin 2006:100–108).

For an excellent treatment of the challenges of creating and maintaining an archive and archive structure, see Trilsbeek and Wittenburg (2006). For additional information on keeping records and on getting funding for archiving, see Pearson (n.d.), which also discusses when to transfer materials and how to maintain updated agreements with archive hosts in incremental submissions.

8.9 Keeping Track of Finances

If funding for the field trip comes from a granting agency, detailed travel costs, daily expenses, and consultant payment records must be kept. It is best to update these records on a daily basis. Receipts should be kept in a safe place and organized to simplify the reimbursement process at the end of fieldwork. It is usually the case that equipment and other expendable supplies will be bought at the home institution, but receipts will be necessary for big ticket items bought in the field such as airline tickets, voltage stabilizers, lodging, and the like. The sponsoring university or grant-administering institution should be consulted before large amounts are spent on unusual expenses. Before going to the field, it is best to discuss with the office in charge of reimbursements how consultants will be paid; the expenses and procedures agreed upon should be documented in writing.

References

Abbi, Anvita. 2001. *A Manual of Linguistic Field Work and Indian Language Structures.* (Lincom Handbooks in Linguistics 17.) Munich: Lincom Europa.

Austin, Peter K. 2006. Data and language documentation. In *Essentials of Language Documentation*, ed. by Jost Gippert, Nikolaus Himmelmann and Ulrike Mosel, 87–112. Berlin and New York: Mouton de Gruyter.

Boersma, Paul and David Weenink. 2010. Praat: doing phonetics by computer [Computer program]. Version 5.1.31, retrieved 4 April 2010. Online: http://www.praat.org/.

Bowern, Claire. 2008. *Linguistic Fieldwork. A Practical Guide*. New York: Palgrave MacMillan.

Branigan, Holly. 2007. Syntactic Priming. *Language and Linguistics Compass* 1/1–2 , 1–16.

Butler, Lynnika and Heather van Volkinburg. 2007. Review of Fieldworks Language Explorer (FLEx) from SIL International. *Language Documentation and Conservation* Vol. 1, No. 1 June 2007, 100–106. Online: http://nflrc.hawaii.edu/ldc/.

Chelliah, Shobhana L. 2001. The role of text collection and elicitation in linguistic fieldwork. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 152–165. Cambridge: Cambridge University Press.

Chelliah, Shobhana L. 2009. Semantic role to new information in Meithei. In *The Role of Semantics, Pragmatics, and Discourse in the Development of Case*, ed. by Jóhanna Barðdal and Shobhana L. Chelliah, 377–400. Amsterdam and Philadelphia: John Benjamins.

Cooper, J. E. 1991. Telling our own stories: the reading and writing of journals and diaries. *Stories Lives Tell: Narrative and Dialogue in Education*, ed. by C. Witherell and N. Noddings, 96–112. New York: Teachers College Press.

Crown, Cynthia L. and Stanley Feldstein. 1985. Psychological Correlates of Silence and Sound in Conversational Interaction. In *Perspectives on Silence*, ed. by Deborah Tannen and Muriel Saville-Troike, 31–54. Ablex: Norwood, NJ.

Crowley, Terry. 2007. *Field Linguistics. A Beginner's Guide*. Oxford: Oxford University Press.

"Digitizing and annotating texts and field recordings in the Awetí Project". Published online in the proceedings of the third EMELD conference "Digitizing and Annotating Texts and Field Recordings", together with the LSA Institute, Michigan State University, Lansing, July 2003. Online: http://www.emeld.org/workshop/2003/paper-drude.html

Dixon, R. M. W. 2010. *Basic Linguistic Theory. Volume 1 Methodology*. Oxford: Oxford University Press.

Dubois, S., and B. Horvath 1992. Interviewer's linguistic production and its effect on speaker's descriptive style. *Language Variation and Change* 4(2):125–35.

Good, Jeff. 2002. *A Gentle Introduction to Metadata.* Online: http://www.language-archives.org/documents/gentle-intro.html.

Healey, Alan. 1964. *Handling Unsophisticated Linguistic Informants*. (Linguistic Circle of Canberra Publications, Series A, Occasional Papers No. 2.) Canberra: Linguistics, Department of Anthropology and Sociology, The Australian National University.

Henderson, John. 2008. Capturing Chaos: Rendering Handwritten Language Documents. *Language Documentation and Conservation* 2(2):212–243. Online: http://nflrc.hawaii.edu/ldc/

Himmelmann, Nikolaus P. 2006. Language documentation: What is it and what is it good for? In *Essentials of Language Documentation*, ed. by Jost Gippert, Nikolaus Himmelmann and Ulrike Mosel, 1–30. Berlin and New York: Mouton de Gruyter.

Hopkins, Jill D., and Louanna Furbee. 1991. Indirectness in the interview. *Journal of Linguistic Anthropology* 1(1):63–77.

Innes, Pamela. 2010. Ethical problems in archival research: Beyond accessibility. *Language and Communication* 30.3:198–203.

Kibrik, Aleksandr E. 1977. *The Methodology of Field Investigations in Linguistics. (Setting up the Problem.)* (Janua Linguarum, Series Minor, 142.) The Hague: Mouton.

Kim, Youjin and Kim McDonough. 2008. Learners' production of passives during syntactic priming activities. *Applied Linguistics* 29(1):149–154.

Ladefoged, Peter. 2003. *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*. Malden, MA: Blackwell.

Matthewson, Lisa. 2004. On the methodology of semantic fieldwork. *International Journal of American Linguistics* 70(4):369–415.

Mc Laughlin, Fiona, and Thierno Seydou Sall. 2001. The give and take of fieldwork: noun classes and other concerns in Fatick, Senegal. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 189–210. Cambridge: Cambridge University Press.

Milroy, Lesley. 1987. *Observing and Analysing Natural Language. A Critical Account of Sociolinguistic method*. (Language in Society 12.) Oxford: Basil Blackwell.

Mithun, Marianne. 2001. Who shapes the record: the speaker and the linguist. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 34–54. Cambridge: Cambridge University Press.

Nathan, David. 2006. Thick interfaces: Mobilizing language documentation with multimedia. In *Essentials of Language Documentation*, ed. by Jost Gippert, Nikolaus Himmelmann and Ulrike Mosel, 363–379. Berlin and New York: Mouton de Gruyter.

Nida, Eugene A. 1950. Training the translation helper. *The Bible Translator* 1:56–62.

Paul, Benjamin D. 1953. Interview techniques and field relationships. In *Anthropology Today: Selections*, ed. by Sol Tax. Chicago, IL: University of Chicago Press.

Pearson, Ruth. n.d. Organizing and Transferring Records. Council for the Preservation of Anthropological Records. Online: http://www.nmnh.si.edu/naa/copar/bulletin5.htm

Samarin, William J. 1967. *Field Linguistics. A Guide to Linguistic Field Work*. New York: Holt, Rinehart & Winston.

Samarin, William J. 1991. Intersubjective and interdialectal variation in gbeya ideophones. *Journal of Linguistic Anthropology* 1(1):52–62.

Sutton, Peter, and Michael Walsh. 1979. *Revised Linguistic Fieldwork Manual for Australia*. Canberra: Australian Institute of Aboriginal Studies.

Tannen, Deborah. 2001. *You Just Don't Understand: Women and Men in Conversation.* New York: Quill.

Thieberger, Nicholas. 2005. The PARADISEC Archive. *Language Archives Newsletter* (5):6–7. Online: http://www.mpi.nl/LAN/

Trilsbeek, Paul, and Peter Wittenburg. 2006. Archiving challenges. In *Essentials of Language Documentation*, ed. by Jost Gippert, Nikolaus Himmelmann and Ulrike Mosel, 311–335. Berlin and New York: Mouton de Gruyter.

Vaux, Bert, and Justin Cooper. 1999. *Introduction to Linguistic Field Methods*. (Lincom Course books in Linguistics 1.) Munich: Lincom Europa.

Whalen, D.H. 1981. The native speaker and indeterminacy. *A Festschrift for Native Speaker,* ed. by Florian Coulmas, 263–277. The Hague, Paris and New York: Mouton.

Woodbury, Anthony C. 2003. Defining documentary linguistics. In *Language Documentation and Description*. Volume 1, ed. by Peter K. Austin, 35–51. London: School of Oriental and African Studies.

Chapter 9
Lexicography in Fieldwork

9.1 Introduction

Within descriptive linguistic fieldwork, as we define it (Chapter 2), lexicography or dictionary writing can be considered a tangential activity. This might seem surprising in view of the modern fieldworker's adoption of the Boasian trilogy, which considers that, for a language to be documented or described,[1] we need a grammar, a collection of texts, and a dictionary (Evans and Dench 2006:10, 12; Dixon 2007). These three components would typically be compiled from fieldworkers' notes, wordlists, grammatical elicitations, and texts.

It is remarkable, then, that Franz Boas himself collected a monumental number of texts (see Chapter 3) and wrote several major grammars, many grammatical sketches, and several short vocabularies, but never published anything that looks like a dictionary. Edward Sapir, certainly a believer in the Boasian trilogy, only published one piece of lexicography in his Southern Paiute volumes (1931) that is long enough to be called a dictionary. Leonard Bloomfield, who also would have been sympathetic to Boas' ideas, did not publish a single dictionary. It would appear that the push towards dictionaries as a necessary part of the trilogy was more the work of Sapir's students, particularly Mary Haas (p.c.). Part of this is certainly Haas' interest in historical and comparative linguistics, which of course necessitated easily comparable lexical compilations.

Why are there so few dictionaries compiled by the early fieldworkers in the Boasian tradition? Boas, Sapir, and Bloomfield must have thought that there were more pressing things to do, such as grammars and text collections. It is not impossible for a linguist to compile a dictionary on the basis of a grammar, text collection, and unpublished lexical field notes, but it is impossible to compile a comprehensive grammar on the basis of a dictionary and text collection. Therefore, early fieldworkers may have seen merit in leaving the time-consuming task of dictionary compilation to a later stage of their work or to others who specialize in

[1]Boas did not draw a neat distinction between documentation and description.

S.L. Chelliah and W.J. de Reuse, *Handbook of Descriptive Linguistic Fieldwork*, DOI 10.1007/978-90-481-9026-3_9, © Springer Science+Business Media B.V. 2011

the task. Since compiling a full dictionary of a language can take several years, the fieldworker typically has time only for shorter dictionaries of the sort compiled by Sapir and advocated by Haas. An example of such a dictionary is Sapir and Swadesh's posthumous (1960) dictionary, which was edited by Haas and which contains about 3,000 entries; by lexicographical standards, this is too short to be considered a full dictionary.

Traditional lexicography is arguably a distinct endeavor from lexicography based on fieldwork. Traditional lexicography is not fieldwork: it is based on earlier dictionaries or on a massive corpus, relies on collectors specializing in spotting "new words", and involves a team of specialists speaking the target language. References on traditional lexicography include the classic manual by Zgusta (1971), and more recent works such as Béjoint (2000), Hausmann et al. (1990-1991), Landau (2001), van Sterkenburg (2003), and Svensén (2009). But all of these will be of only limited use to the fieldworker.

Instead the fieldworker should look for advice in manuals or essays written specifically for fieldworkers dealing with endangered languages or languages without a written tradition. The most useful are: Bartholomew and Schoenhals (1983), Crowley (2007:38-44), Frawley et al. (2002), Haas (1962), Haviland (2006), Matteson (1976), Mosel (2004), Nichols and Sprouse (2003), and Stebbins (2003). Another useful resource is the Bibliography of Lexicography (Hartmann 2007).

In this chapter, we distill from these works some advice for the fieldworker on collecting words which can be used for wordlists and for lexica or dictionaries. A wordlist is a carefully transcribed and translated list of words which may be designed for a specific analytic purpose, such as the acoustic study of vowel quality or the morphological and phonological shape of compound nouns. We use the term lexicon to refer to a smaller version of a dictionary, containing much the same information; thus we use the terms lexicon and dictionary more or less interchangeably.

9.2 Wordlist Elicitation

In this section, we provide a methodology for creating a wordlist, an essential document for descriptive fieldwork. As one of the first interactions between fieldworker and native speaker in a formal interview setting, the wordlist gives the fieldworker the opportunity to set a positive tone at the beginning of their working relationship. Since eliciting words is a relatively easy task for the speaker, the fieldworker can share control of the session and allow for task-related conversations with topics determined by the speaker. The task of wordlist elicitation must be carefully executed, since it forms the basis of phonetic and phonological analysis; furthermore, since some of the words may be morphologically complex forms, morphological structure can be determined from wordlists as well. Collecting wordlists from several languages spoken in the same area allows for cross-linguistic comparison and possibly for determination of genetic relatedness. Recording words of specific semantic domains also documents material and intellectual culture.

9.2.1 Using a Basic Wordlist

To prepare a schedule (or plan) for wordlist elicitation, the fieldworker should first determine if any previous wordlists have been collected for the language under study or for related languages. The second step is to re-elicit or get equivalents for 200-500 words from these earlier lists.[2] If an existing list for the relevant area does not exist, the fieldworker could modify the 200 wordlist proposed by Morris Swadesh as a way to compare languages to see how closely they are related (Swadesh 1955). This list is available online at http://lmo.wikipedia.org/wiki/Swadesh_list. In either case, only culturally appropriate words and locally relevant concepts should be included. A note of warning about re-checking previously elicited lists: a language can change even in the lifetime of a consultant. For example, Alec Coupe (p.c.) tells us that he worked with a language that had five lexical tones. However, according to descriptions based on work with the same consultant 30 years earlier, the language had six lexical tones, so assuming there was no error in analysis at either point in time, the language appears to have changed.

9.2.1.1 Picture Prompts

Not all speakers will be able or willing to work with drawings, printed pictures, or photographs – for example, speakers from rural areas who are not accustomed to working with printed materials. This is because the speaker might think that what is required is the name of a particular token of an item rather than the general term for that item, e.g. when shown a picture of a horse they may want to give the name of a particular horse or type of horse rather than the general term for 'horse'. In addition, speakers may not share our conception of a picture as a representation of something or of some event which exists outside of the representation itself.

However, speakers from more urban areas may be able to use visual prompts for word production. Thus the fieldworker could try a name-the-picture task for eliciting nouns, verbs, adjectives (including emotions), prepositions and adverbs. A picture-based schedule should be comprised of a carefully selected set of pictures organized by theme and presented in a neat format to the consultant; this will thus take more time to prepare than a wordlist. The fieldworker should take time to frame the elicitation question so that it is clear what information is needed. If, for example, the speaker is shown a picture of a man in a boat, fishing, what is it in that picture that

[2]The number of words suggested for initial phonetic fieldwork varies. Some say 200, others – like Grimes (1995: 9) –say 300.

the speaker should identify? The fieldworker has to be concerned about correctly interpreting speaker responses. Is the speaker providing a word for the intended item, or for something else? The fieldworker will need to know a little bit of the language to assess if the speaker has provided the required response since, depending on how the question was phrased, it could be any one of the following: the word for 'boat'; the occupation of fishing; the act of fishing on a habitual basis; the physical environment (e.g., the lake the boat is in, or the mountains surrounding the lake), the emotional state of the fisherman, etc.

Needless to say, the pictures used should be culturally appropriate and locally relevant; pictures of women in miniskirts would not be well received in conservative Sri Lanka (Canagaraj 1999), and as we found out, speakers of Tangkhul Naga (a Tibeto–Burman language of Manipur, India) are uninterested in translating books with illustrations of dragons and damsels in distress. Similarly, Mayer's *Frog, Where are You* (1969) or the *Pear Story* (Chafe 1980), even though used quite often, could be too culturally specific to elicit predictable terms. For example, even English speakers from India will identity the pear as a mango or guava because pears are not commonly found in India.

Keeping this in mind, the fieldworker may get pictures from a photograph website such as http://www.flickr.com, or from dictionaries, foreign language teaching manuals, or books about and with photographs of the flora and fauna of the relevant region. Crowley (2007:108) and Coelho (2006:4) warn of the limitations of using pictures from books. Speakers may not recognize individual species if these are seen out of their natural environment. They may also need to perceive a specific quality of the tree – the bark or the smell, for example – in order to identify the species, and this may not be part of the illustration. Just because botanists have similar names or cover-terms for related species – *oak* for types of *oak*, for example – does not mean that the target language will lexically connect related species. Crowley (2007:108) also notes that the same tree may have different names at different stages of growth.

For names of local species, it is best to have the actual fruit or flower in hand so that the consultant can study the object before naming it. To accomplish this, Coelho (2006:4) trekked with her consultant – and sometimes a botanist as well – through the Mudumalai Wildlife Sanctuary (South India) to elicit plant names. For accurate plant classification, she thought like a botanist and took photographs of plants from five perspectives (overall, stalk or branch with leaves, flowers, fruit, and other salient characteristics).

While animals are harder to locate in the outdoors, actually seeing the animal is sometimes necessary when command of the contact language is weak. Here is an example. The Lamkang word *theipa* was glossed as both 'rabbit' and 'squirrel' in our texts and we had a hilarious time looking for the correct gloss. The rabbit is fast. Well, so is the squirrel. The rabbit has protruding incisors. Well, so does the squirrel.[3] Finally, one of the speakers went outside and pointed to a squirrel, which fortunately are abundant in the part of the United States we were in at the time. We had identified *theipa*.

[3]It is interesting that none of the speakers thought of identifying a 'rabbit' by its long ears and short tail, and a squirrel by its short ears and long bushy tail, as English speakers presumably would do.

The fieldworker will also have to bear in mind that an animal may have different names depending on whether it is domestic or wild, or depending on its pattern of colors or spots.

Maps are good for re-eliciting and talking about place names (Tobias n.d.), assuming the native speaker is familiar with the concept of maps. Nancy Caplow (p.c.) points out that trekking guides are often familiar with maps even if broader community isn't.

It may not be possible to use photographs of people from the region; the field-worker would need to know how community members feel about having their pictures taken and used in this manner.

A good source for picture-driven elicitation is Bouquiaux et al. (1992), which includes an extensive collection of line drawings for eliciting basic actions and technologies in an African village context, including forges, weaving, basketry, architecture, fishing, trapping, and body ornamentation. The book also has drawings regarding the human body and anatomy, African animals, and African plants. Our only reservation would be that some of the pictures of human anatomical deformities and infirmities might make some consultants uncomfortable, but we have not used these pictures for elicitation ourselves.

More discussion of elicitation with pictures and video prompts is in Section 12.2.2.2.

9.2.1.2 Compiling an Expanded Wordlist

Extended culturally-specific wordlists for the relevant language family or subgroup can be used to as elicitation schedules. The resulting wordlists are useful for cross-language comparison. A sampling of such wordlists is provided in Table 9.1. Further examples are listed in Table 12.1.

9.2.1.3 Texts

A project's wordlist continues to expand during and after the field trip. Elicited sentences will yield new words. Words may be added because they are heard in conversation outside a formal field session. Healey (1975) suggests that all new words be noted, even when the meaning is not known at the time – translations can be obtained at a later date. In fact, it is often easier to elicit the meaning of a target language word than to elicit a translation of a contact language word. That is why texts are such great sources of words: along with native speaker input, they provide the context needed to make an accurate translation. See Sutton and Walsh (1979:18) and Kibrik (1977:63-67) for further discussion.

There is a great danger in eliciting words without context. Ken Hale recalls that a native speaker of Alyawarr was incorrectly identified as someone who knew only 81 words of the language, while Hale was sure that he must know many more. How did this incorrect characterization of a native speaker take place? Hale hypothesizes

Table 9.1 A sampling of general lexical elicitation schedules

Reference (chronological)	Language family/ area covered	Pages where suggestions for use of the schedule are given	Lexical items to be translated	Comments
Pittman 1948	General	133	708	Useful in that it is organized from least abstract to more abstract.
Bee and Pence 1962	Papua New Guinea	64–65	153 Words; 37 phrases	For cross-language comparison. The first 50 words can be elicited through gesture. Organized for determining language relatedness.
Capell 1965	Australia	1–9	446	For the investigator with no prior experience in fieldwork. Superseded by Sutton and Walsh (1987).
Kibrik 1977	Primarily the former Soviet Union	35–39	Approx 1,482	Appendix 4, pp. 99-110 is a general list of sentences for the first encounter with the lexicon and grammar of the target language. Appendix 5, pp. 103-124 is a short dictionary, thematically organized.
Matisoff 1978	Southeast Asia		200	Intended to make the Swadesh list culturally appropriate by replacing items like *snow* with items like *monkey*, *banana*, and *bamboo*.
Bartholomew and Schoenhals 1983	Primarily for Mexico and Latin America	293–294	About 2,200	pp. 295–338 is a basic vocabulary of Mexican rural Spanish, with English glosses.
Sutton and Walsh 1987	Australia	4–6	2,195	Structured by semantic field and lexical category.
Bouquiaux et al. 1992. Questionnaire 2	Primarily for Africa	183–185	221	Words in sample sentences.

Bouquiaux et al. 1992. Questionnaire 3	Africa	45	704	Combines Greenberg's *List of Words for African Languages* (items 1–547) and A. E. Meeussen's *Supplement* (items 548-647).
Abbi 2001	South Asia	260		Reduplication and explicator verbs.
Matisoff 2003	Southeast Asia		74 Words on kinship and people, 342 on plants 343, 154 on natural objects; 233 on animals, 364 on body parts	Organized by semantic field. Available online.[4]
Gil and Shen 2004	General	Instructions are interspersed with the list	65	Intended to elicit metaphorical expressions. Available online.[5]
Roberts and Keith Snider (2005-2006)	Africa	4–5	1,700	For cross language comparison. Items in wordlist are given in French and English. Semantically classified. Available online.[6]
Dixon and Blake (1979) 1981, 1983, 1991, 2000)	Australia	none	Between 1,000 and 2,000	A five-volume series of language sketches, each with a vocabulary list organized by semantic field.

[4] http://stedt.berkeley.edu/html/questionnaires.html
[5] http://lingweb.eva.mpg.de/jakarta/docs/English.pdf
[6] http://www.sil.org/silewp/2006/silewp2006-005.pdf

that the investigators must have tried to get a list by translating from an alphabetic list of words. When this is done, "you get tired after a while. You can't remember … you try with English, you sit down and try to remember…You might start with *ahert* [bilby] … then you might get *aherr* [kangaroo]…OK what's the next word we'll put there – can't think of one … You might get to, "I can't think of any more words – maybe that's it." (Green 2001:36)

At an early stage of investigation, the fieldworker can draw out particular lexical items without directly eliciting them. For example, the native speaker could be asked to talk about his or her family. This would give proper names and kinship terms which could then be used as a springboard for further discussion. If a list of verbs is needed, the native speaker could be asked to describe what he or she did yesterday. See Sections 12.2.2.8, 12.4.4 and 13.4 for other advice on working with texts.

9.2.1.4 Dialect Surveys

At later stages, the fieldworker will want to get less common words, such as vocabulary used in rituals or other traditional activities, archaic words, and words known to be specific to a particular region. Vaux and Cooper (1999:158) recommend eliciting such vocabulary through the use of dialect surveys which sometimes are designed to look for less common words. Examples are Redard (1960), Guillaume (1963-1966), and Wolff (1969). Other examples of such schedules are the wordlists used in the Linguistic Atlas of the Gulf States.[7]

9.2.2 Organizing Lexical Collections

Several approaches may be useful in organizing and presenting the lexical items once they are collected:

- Alphabetically or reverse-alphabetically, by target language
- Alphabetically by gloss
- Thematically or by semantic field
- By groups of synonyms, antonyms, and homonyms
- Alphabetically by stem or root
- By lexical category

However, while those may be useful ways to organize data in the final product, they should not dictate how lexical data is collected. For instance, elicitation schedules should not simply be alphabetized lists of words; it is easier for a speaker to recall lexical items when elicitation occurs within a semantic field. Likewise, eliciting

[7] http://www.proquest.com/en-US/catalogs/collections/detail/Linguistic-Atlas-of-the-Gulf-States-333.shtml

around a specific cultural event or theme can prompt the speaker to offer related items which the fieldworker would not have thought to ask about (Kibrik 1977:35-39). To elicit nouns, for example, a schedule should group together nouns in the same semantic field. Some examples of semantic fields taken from Healey (1964), Abbi (2001), Sutton and Walsh (1979), and the Stanford Linguistics Fieldwork Committee (2006) are:

- Social relationships: Kinship and address terms, professions, personal names. Note that kinship terms can be tricky to collect because the target language may not have the same relationships coded in the same way as the contact language.
- Items of material culture: Furniture, architecture, clothing, jewelry, cooking utensils, spices (which Abbi (2001) says are useful for getting compound nouns in Indian languages), types of food, toys, sports, and sporting equipment.
- Surrounding environment: Seasons, weather, time, cosmological terms, geographical or geological terms, flora and fauna. Elicitation of folk taxonomies, such as ethnobiological (ethnobotanical and ethnozoological) taxonomies, is a time-consuming and specialized enterprise. It is not the sort of thing that the fieldworker will start with in word gathering, but if the fieldworker plans on writing a dictionary, this terminology and related taxonomies will have to be elicited. Additional advice regarding ethnobiological taxonomies and terms is given after this list.
- Physical: Body parts, ailments, cures, medicines. There may be certain sensitive or secret elements in the discussion of body parts which can be skipped if the consultant is not comfortable providing them (see Section 9.2.3).
- Numerals.
- Emotions.
- Mental and speaking activities: Remembering, forgetting, thinking, paying attention to, requesting, thanking, talking, telling.
- Religious terms: Deities and spirits, temples, sacred sites and pilgrimage destinations; festivals, clothing and food for feast days, religious ceremonies such as puberty, naming or marriage ceremonies. There may be certain sensitive or secret elements in the discussion of religious terms which can be skipped if the consultant is not comfortable providing them (see Section 9.2.3).

More specific advice regarding folk taxonomies and ethnobiology is in Conklin (1962), Berlin (1992), and Fleck (2007). Sutton and Walsh (1979:18) caution that natural species names are hard to elicit. It might be necessary to work with a specialist, matching Latin names with pictures or taking pictures during a walking tour and matching these to Latin names later. See Crowley (2007:108) who writes on the difficulty of accessing native botanical knowledge out of context. Wherever possible, translations should be visually confirmed, because varieties of the contact language may collapse distinctions; for instance, in Indian English, the magnolia tree and the mimosa tree can both be referred to as a "magnolia".

An early exemplary treatment of the ethnozoology of Tewa, a language of New Mexico, is Henderson and Harrington (1914). A modern Tzotzil ethnobotany is Breedlove and Laughlin (1993). This is actually a thematic ethnography extending beyond ethnobotany, since it also deals with things such as fences and houses.

Eliciting names of common fauna can at the same time be a rich source of syntactic information. For example, Dixon (2010:297) points out that, having elicited the word for a type of bird, the fieldworker might ask, "What does that bird do?", to which a consultant may well provide a description of habitual action, possibly with verbs of movement, e.g. "This pigeon eats the flowers of that palm tree and lives at the top of it."

Other schedules based on thematic organization and semantic fields can be developed for verbs, e.g. verbs of handling, movement, destroying, and becoming. One example is in Kibrik (1977:90-96), and a comprehensive source for compiling these lists is Levin (1993). It is usually best to elicit verbs in simple sentences, using picture prompts where possible. See Sections 12.3.1 and 12.3.2.1 for more discussion of paradigm elicitation.

Adverbs are a little tricky to elicit without context, so this is a list that should be derived from texts (Stanford Linguistics Fieldwork Committee 2006) and then fleshed out by eliciting antonyms. An interesting schedule for eliciting verbal, adverbial, or adpositional spatial meanings is in Kibrik (1977:83-88).

Adjectives can more easily be elicited in semantic fields: human attributes ("tall, short, fat"), consistency ("flexible, hard, soft"), colors ("black, blue, green"), tastes ("sweet, spicy, salty"), emotions ("sad, happy"), and states of being ("lazy, growing, ill"). It is possible that, in the target language, these qualities will be expressed not by adjectives but by stative verbs (translatable as 'be tall, be soft, be blue, be sweet', etc.).

It is always difficult to distinguish between grammar and lexicon, especially in the case of idioms and collocations, but also acutely so in polysynthetic languages which include lexical or productive derivational material within words (Evans and Dench 2006:12-14; Mosel 2006a:46-47). This is an area where the fieldworker's notes will benefit from a reasonable amount of redundancy and cross-referencing.

Although organizing entries by semantic field rather than alphabetically is useful in elicitation, and pedagogically sound, there is no generally agreed-upon system for organizing a thematic dictionary. The fieldworker and reader will be able to find items faster if an alphabetical index is added to a thematic dictionary. An unusually large thematic dictionary is Kari (2007) for Dena'ina, an Athabascan language of Alaska.

9.2.3 *Methods for Word Elicitation, Recording, Cataloging, and Checking*

To create a wordlist in the target language, the fieldworker must interview speakers, and record and catalog words. Techniques for these processes are reviewed below:

9.2.3.1 Interviewing Techniques

During the interview, the fieldworker concentrates on transcribing and on correlating translations to elicited words. Sample sentences should also be elicited.

The importance of illustrative sentences, ideally taken from one's fieldnotes or texts, or elicited for the entry, is discussed in Bartholomew and Schoenhals (1983:59-69).

A typical interview involves the fieldworker asking the native speaker for equivalents in the target language. Here are some details of what such an interview might look like:

> Fieldworker: *What is this? How do you say…? What is the word for…?* Depending on whether the prompt is a picture, a word in the contact language, or an action that needs to be translated, one of these questions will be asked of the native speaker. The fieldworker must look for confirmation that the native speaker has accurately understood what needs to be translated: for example, does the native speaker recognize the word in the contact language? Are they looking at the right part of the picture that needs to be translated?
>
> Native speaker: *The word you are looking for is*__________.
>
> The native speaker provides the word, usually saying it only once until they catch on that the fieldworker needs the word to be repeated in order to transcribe it.
>
> Fieldworker: *So the word is* __________.
>
> The fieldworker says the word back to make sure all sounds have been correctly heard. This is especially important in a noisy environment or with a soft-spoken native speaker. The word may have to be repeated several times. This can get awkward, as the native speaker may grow annoyed or bored by repeated requests for repetition. Some speakers may not be able to repeat exactly the same form again, especially if the word is morphologically complex. Some speakers will slow down speech with each repetition of a word so that vowel quality, intonation, stress, and tone are affected. Thus there should be a limit to the number of repetitions required of the speaker; Foley (2002:133) recommends limiting repetition requests to five; Edward Sapir told his student Li Fang Kuei that one should not ask for more than three repetitions (Randy LaPolla p. c.). The fieldworker could try to figure out what is wrong: the vowel? or perhaps the tone? If there is still some confusion, that particular word can be returned to at a later stage.
>
> Native speaker: *That's right. You said it right, you sound like a native.*
>
> It is good to remember that some native speakers will accept a pronunciation if it comes anywhere close to what they said. If the native speaker is too accommodating, additional techniques will have to be used to check transcriptions, such as this one described by Foley (2002:133): purposely repeat the word with one sound off and note how the native speaker reacts. Hopefully, he or she will repeat the correct form. Another possibility is to provide several alternant pronunciations of a word and ask which one sounds best.
>
> Fieldworker: *Okay, let me write that down…so, the word for ___ is ____.*
>
> The fieldworker writes down the elicited word and says it back with the translation. If the referent of a term is still unclear, the fieldworker can confirm the translation at this point. If, for example, the word being elicited is "blue", the fieldworker could show blue and green color chips to see if there is difference in terms for the two colors.

Many fieldworkers report that they record word-list elicitation sessions in compressed format (mp3, for example) so that they have a record of what was discussed, and they can use this to check their notes later on. As discussed in Section 8.3.3, wav files are required for accurate instrumental study.

The precise questions used during word elicitation sometimes have to be carefully thought out. The fieldworker should think of how he or she would define or further describe a word or picture prompt. For example, if the fieldworker asks for a translation of ‘hen’ but the speaker does not know this word, the fieldworker can then ask for ‘male chicken’.

The attempt to find accurate equivalents for words in the target and contact languages will be refined as more tokens of the target language item are seen in context. For Healey (1964), it takes at least 10-20 tokens of a stem and 50-100 tokens of a function morpheme, in context, to fully understand the form’s meaning and use.

Fieldworkers should keep in mind that the distinctions made in the contact language may not be found in the target language. For example, in many languages the same word is used for “blue” and “green”; “nail” and “finger”; or “milk” and “breast”. This does not, of course, indicate a gap or shortcoming in the target language, and fieldworkers should not show surprise or disdain when coming upon such lexical patterns.

Dixon (2010:325) suggests taking some simple props to the field to assist with getting at the meanings of words: for example, a piece of string can be used to determine what the consultant means when he says ‘pull’. Does the word refer to a single motion of pulling, or to repeated tugs? Dixon (2010:292) also discusses contrast as another useful method in lexical elicitation. Color terms and adjectives for example, can be more accurately glossed when elicited in contrast.

The main principles for elicitation of meaning in lexicography are given in Vaux et al. (2007:77-86). Useful ideas can also be found in Bartholomew and Schoenhals (1983:70-92), Mathiot (1967), Beekman (1975), Cruse (1986), Nida (1958), and Haviland (2006:134-151).

When working with different dialects, it is always interesting to ask if there is an alternate word for an elicited item, i.e. “Is there another way of saying X?” This question might elicit true variants, but the speaker might also produce variants that suggest a dialect difference where none exists.

Ethnographic information should be included as annotation to word collection (Thompson 1989). This makes sense when one considers how translation introduces the possibility for error. As discussed in Franchetto (2006:186-187), the fieldworker must be careful in translating words so that the whole range of meaning is covered. She gives the following example: in Kuikuro (a Carib language of Brazil) *tolo* means ‘bird’ but it can also mean ‘pet’, ‘a song genre’, or ‘my lover’. It is also not uncommon, in languages of the world, for the word “bird” to refer to ‘penis’ in some contexts, so caution has to be advised sometimes with the most innocent sounding words.

Together with ethnographic information comes cultural sensitivity. It is counterproductive to elicit taboo forms or to force consultants to give translations when they are not comfortable doing so. If they are interested in sharing forms, they will. Sometimes it just takes a few years or a certain type of consultant to contribute taboo terms. Speakers may use euphemistic substitutes. Crowley (2007:105) suggests skipping from navel to knee when first eliciting body parts to avoid embarrassing the speaker, waiting to get the rest after a secure relationship has developed between fieldworker and consultant. Taboo forms may be sexual or scatological, or they may be about menstruation, but they may also involve place names, relatives to be avoided because of a mother-in-law taboo, names of people no longer living, names of gods or mythological beings, spirits, witchcraft, or dangerous animals such as bears, big cats, crocodiles, sharks, snakes, and wolves.

In some cases of language obsolescence, it has been reported that word elicitation cannot occur in a straightforward manner because words themselves have become sacred or precious possessions to be displayed only on special occasions. Moore (1988) reports, for example, that for Wasco (Chinook) speakers, "the standard elicitation-interview protocol constitutes one appropriate event-type, but only for some speakers, and for them only some of the time" (Moore 1988:463). One younger speaker thought of words as having special mythical qualities, treating them as myths by only 'reciting' them in the wintertime, the prescribed season for myth-telling in many Native American cultures. Also, Moore reports that speakers may insist that an object be present before it can be named. While conversing in English, if the speakers refer to an object, then the Wasco word for that object might be recalled and uttered (465).

Another type of word that cannot be elicited for the Wasco–Wishram are personal names (Moore 1988:464-465). Names bestowed through naming ceremonies are not to be referenced outside of the naming ceremony (465).

To summarize, a word-elicitation session will be most successful if the fieldworker:

- Is fully prepared for each session, with an organized elicitation schedule and well-thought out prompts and phrases requesting translation.
- Demonstrates to the native speaker that each session has a goal, and the fieldworker knows how to reach that goal.
- Takes organized notes. This will reassure the native speaker that the fieldworker is serious about the task at hand and so the speaker's time is well spent in this activity.
- Transcribes accurately so that he/she can practice and produce language forms correctly. This will also demonstrate the fieldworker's sincerity in the documentation process.
- Encourages literate native speakers to participate in translating lexical items.

9.2.3.2 Recording Techniques

The fieldworker also needs to think about how to record word elicitation sessions. It is useful to think of recording as occurring in three stages. In the first stage, recording occurs alongside transcription and translation. When working with an endangered language, it is advisable to keep the recorder running through the whole word elicitation session, since it is possible that this will be the only chance the fieldworker will have to record these words.

In the second stage, the fieldworker prepares wordlists for instrumental study that will lead to phonetic and phonological analysis. The first task at this stage is to compile the information gathered during elicitation and come up with organized wordlists such that the source of each word can be identified: which notebook the word is taken from, the page number, the unique item number, the target language form, the contact language gloss, and, if available, the spelling in the practical orthography.

The fieldworker is now ready to record the wordlists. The easiest scenario is one where the native speaker can read the practical orthography and contact language. In this case, the speaker should say the number of the word in the list, say the gloss in the contact language or lingua franca, and repeat the word twice in the target language (Grimes 1995:10). There should be a pause between tokens. Alternatively, the fieldworker may read the contact language prompt and item identifier information, and then ask the native speaker to say the word in the target language. Issues such as reading pronunciation and artificially slow speech should be kept in mind; the reading of wordlists should always be balanced with production of words in frames. See Section 10.2 on recording words for instrumental analysis.

If it is not possible to divide word elicitation between several sessions, then both elicitation and recording for instrumental analysis will have to done at the same time. Sutton and Walsh (1979:7) suggest that the fieldworker first write down all speaker responses, and then re-elicit in list form with the recorder running. The transcriptions can be checked later, during re-elicitation. To keep the whole process manageable, the fieldworker could re-elicit and record in chunks of 50 words at a time (Abbi 2001). While some native speakers like this sort of task because they see exactly what is being accomplished and appreciate their role in reaching that goal, other native speakers may tire or bore easily. As mentioned before, the field-worker must be aware of the energy and interest level of each consultant he or she works with, and modify sessions accordingly. Also, speakers must be carefully selected for this task since some are not able to read from a list naturally. See section 10.2 on the problem of list intonation.

In the third stage, the fieldworker checks transcriptions against the recording and prepares an elicitation schedule – a plan for further elicitation – to clear up any discrepancies between the recordings and transcriptions. During this checking stage, the fieldworker could play back the confusing segments to the native speaker and then record their comments on a second recorder. To help evaluate material from one speaker, transcriptions can be checked with a different speaker, if possible.

A note of caution: the initial work of word elicitation is time consuming (although not nearly as time consuming as text transcription). It can also be frustrating, depending on the phonological complexity of the target language; a tone language, for example, can throw the fieldworker into a panic, as each attempt at transcription seems to require revision (see discussion in Hyman 2001). The fieldworker might find it useful to come back to and recheck a list after becoming better attuned to the target language.

If the fieldworker finds considerable variation in the pronunciation of different speakers, manifested as inconsistencies in his/her transcription, it might be helpful to obtain what Ladefoged (2003) calls a “community pronunciation”. The way to do this is to record wordlists in groups, with a well-respected speaker acting as leader. This [usually older] speaker would say the word first, and others would individually repeat the word or comment on it. Crowley (2007:107) also finds 'brainstorming' in groups helpful in coming up with words, and Holton (2009:170-171) reports on the positive experience of attracting community members to contribute words in their area of technical expertise in large group sessions that turned out to be social events. However, Steve Morey (p. c.) found eliciting words from a group to be time consuming because of all the discussion and digressions; so sometimes it is better to start list elicitation with one person.

Models of modern fieldwork-based lexicography include: Badten et al. (2008) for Siberian Yupik Eskimo; Day (1994, 1995) for Abenaki (Algonquian); Enrico (2005) for Haida of Alaska and British Columbia; Hopi Dictionary Project (1998) for the Hopi of Arizona; Laughlin (1975) for a dialect of Tzotzil (Mayan); Matisoff (1988) for Lahu, a Lolo–Burmese language; Parks and Pratt (2008) for Pawnee (Caddoan); Sylestine et al. (1993) for Alabama (Muskogean); and Ullrich (2008) for Lakota (Siouan). Of these, the Hopi dictionary and the Tzotzil dictionary (both with about 30,000 entries)[8] are the most impressive in coverage.

Exemplary dictionaries which combine philological work with fieldwork products are: Jetté and Jones (2000) for Koyukon Athabascan of Alaska, and Karttunen (1983) for Nahuatl.[9]

The works listed here benefited from extensive collaboration with native consultants, and are exemplary in the richness of grammatical detail provided, as well as – for most – richness of ethnographic detail. They also used computational aids to various extents.

9.2.3.3 Techniques for Cataloging Lexical Data

Data from word-elicitation session notes must be organized so it can be easily reordered (by part of speech, or alphabetically by gloss, for instance) and retrieved (to a printable file, for instance). As discussed in Kibrik (1977:63-67), Healey (1975), Grimes (1995), and Abbi (2001), information on each word collected should allow for easy reclassification by different fields and for cross-indexing of information. The best approach seems to be to have a template that allows for a maximum amount of information that can be filled in gradually. The information should be organized from basic to more esoteric; fields that do not get filled in could be deleted later. Useful fields are:

1. Phonetic transcription.
2. Phonemic transcription.

[8] Haviland (2006:133) states, however, that there are over 35,000 entries in Laughlin (1975).

[9] There are also, of course, recent dictionaries based entirely on philology, such as Masthay (1991, 2002), and Laughlin and Haviland (1988) (see Section 5.2).

3. Orthographic representation.
4. Meaning.
5. Source
   a. Speaker identifier.
   b. Notebook numbers and pages were discussed.
   c. Reference for examples from previously recorded data such as a wordlist from a published source.
6. Comments: An additional cell should be provided for added comments.
7. Morpheme type: e.g., affix, stem, root, or free form.
8. Lexical class (part-of-speech).
9. Sub-class membership: e.g., noun class, semantic verb class, transitive/intransitive verb.
10. Semantic classification.
11. Example sentences and/or abbreviated references to examples from texts.
12. Cross-reference to where conjugation or declinations can be found.
13. Related words such as compounds formed with this word; words related through polysemy or homophony.
14. Syntactic and morphological properties (e.g. co-occurrence restrictions or lexical combinability with other morphology).
15. Context of use and other cultural information.

9.2.4 Database Management and Lexicography

Data management software such as TOOLBOX/FLEx have been specifically created for linguists and are handy, among other things, for classifying, grouping, and sorting data, and linking lexical items to sample sentences in text files. However, our interviews of linguistic fieldworkers suggest a pretty even split between those who use prepackaged software and those who design their own template using generic database management software (e.g., Excel, Filemaker). It might be useful to adapt a cataloging system developed by another linguist working on a typologically or genetically related language. This would help in setting up generally appropriate fields which could be modified as the fieldworker's understanding of the target language increases.

The literature on computer-aided lexicography, including dictionary maintenance and presentation tools, is expanding rapidly. Some useful references are Coward and Grimes (1995), Heather and Rossiter (1988), Poser (2002a, b), and Frawley et al. (2002). We do not mention Internet sources since these are easily located and are constantly updated. See Abney et al. (2006) for some useful leads on online resources.

Of course, a lexical database and the resulting lexicon created using the latest technology will not necessarily be analytically superior to one which is created the old-fashioned way (Dixon 2007). We know that it is the fieldwork underlying these works – and not the computational expertise – that makes them exemplary. Just as there are a few poorly designed and poorly researched dictionaries that were made with computational help, there are good dictionaries made without technological support. It is heartening to remember that in even in the pre-computational age, and before the Haas-inspired focus on the dictionary component of the Boasian trilogy, there have been cases of extremely diligent fieldworkers who compiled dictionaries which are fully comparable – in terms of grammatical and ethnographic detail and accuracy – with the dictionaries mentioned in Section 9.2.3.2. A few examples are: Abraham (1964) for Somali, Bridges (1933) for Yahgan (or Yamana) of Tierra del Fuego, Jäschke (1881) for Tibetan, Man (1923) for Andamanese, and de Smedt and Mostaert (1933) for Monguor, a Mongolic language of China.

If we are serious as an academic community about encouraging indigenous scholarship on endangered languages, we must accept that not all community-based researchers will have or will want to have access to a computer, and they may prefer paper file slips. However, because digital files can be archived, shared, used for generating working lexicons, and updated more easily than hard-copy, organizing a lexicon using computer software is preferable to using hard-copy file slips.

Wordlist collection and analysis is never complete (Crowley 2007:111): words are added as more texts are glossed, and as cultural and social information (taboo words, gender-specific, dialectal or archaic forms) is gradually updated. The fieldworker should archive collected materials as much as possible, but should not have illusions about a full dictionary resulting from a few years of fieldwork. Mosel (2004) provides the best summary of what can reasonably be achieved by a fieldworker in less than a year.

9.3 Bilingual Dictionaries

Before making a lexicon or bilingual dictionary available to others, it is advisable to ask a native speaker to re-check transcriptions and translations for accuracy. Bartholomew and Schoenhals (1983:249-250) report on trained literate native speakers who evaluated dictionaries for "contradictions, omissions, ambiguous glosses, unrecognized sense discriminations, [and] inconsistencies in handling similar kinds of information." Native speaker evaluation also helped to show if the organization of the dictionary was intuitive and helpful to the native speaker, and if the glosses of the sample sentences and the sentences themselves were representative of the listed meaning. The native speaker input was not meant to be used for specific revisions but was rather taken as a general statement of how the organization of the dictionary could be improved.

A fieldwork-based dictionary will be bilingual (or even multilingual). The field-worker will of course gloss or define the lexical data in the contact language or in some other language that s/he is most familiar with. However, for a published dictionary, the choice of this contact or other language will need to take into account the wishes of the community. For example, the superb dictionary of Tzotzil by Laughlin (1975) is not easily used by the community because it is

Tzotzil–English, English–Tzotzil, and of course the other language of the Tzotzil community is Spanish (Haviland 2006:133).[10]

A bilingual English (or other world language) to target language dictionary is just as useful as the reverse, i.e., target language to English (or other world language). Actually, in the case where the community has almost entirely shifted to a world language and consists of learners of the target language, the world language to target language section is vastly more useful than the reverse. This is the reason why Parks and Pratt's (2008) dictionary of Pawnee has a first and more expanded English–Pawnee section.

There is no simple way of dealing with languages with complex and irregular prefixation, such as Athabascan or Caddoan, or languages with infixation, such as many Austronesian languages. One can alphabetize by first letter of the word, in which case words that are derivationally or inflectionally related will potentially occur very far from each other. Alternatively, one can alphabetize by the first letter of the stem; but this requires considerable linguistic sophistication on the part of the reader, because s/he must be able to tell where the stem starts, which might not be easy if the stem undergoes extensive morphophonological alteration. One solution is to rely entirely on electronic dictionaries, as advocated by Poser (2002a, b), and Maxwell and Poser (2004). Another solution is to adopt a paper dictionary with a compromise approach. This approach might entail some principled inclusion of inflected and derived words as entries, as well as the inclusion of entries departing from stems, with an extensive system of cross-referencing between both sorts of entries. Bartholomew and Schoenhals (1983:41-49) discuss special problems with agglutinative languages.

Since dictionaries are more visible or salient to the public than are grammars or text collections, community members are more likely to have strong opinions about dictionaries.[11] Speakers may object to dictionaries for religious, puristic, or political reasons. Other speakers will consider the dictionary (particularly a handsome paper copy to be displayed on a bookshelf) to be an emblematic part of their linguistic heritage (Tonya Stebbins, p.c.). Extra care and diplomacy in securing community support and collaboration is crucial. Also, because dictionaries are so salient, it is expected that they be written using spelling conventions the community can adopt and read. Orthography and standardized spelling issues have more to do with language engineering than with descriptive fieldwork, but the fieldworker will not be able to avoid them. See Henne (1991), Hinton (2001), Seifart (2006), and Grenoble and Whaley (2006) for discussion of this point.

These and other problems of dictionary writing for endangered language communities are described by Bartholomew and Schoenhals (1983), based on their experience

[10] This imbalance was corrected in a later dictionary, which is trilingual Tzotzil–English–Spanish (Laughlin and Haviland 1988).

[11] Texts containing religious or otherwise sensitive material can also be objected to, of course, but these are more easily predicted than objections to a dictionary.

in writing indigenous Mexican language dictionaries for the Summer Institute of Linguistics; by Stebbins (2003:121–262), based on her experience with Sm'algyax (Tsimshianic of British Columbia); by Corris et al. (2004) based on experience in Australia; by Mosel (2006b), mostly based on experience in Samoa and Papua New Guinea; and most cogently by Crowley (2007:35-44), based on experience in Vanuatu and Australia. Dictionary development for indigenous communities on the basis of older materials is discussed by Rudes (2002). (See also Section 5.2.)

Increasingly, dictionary creation is seen as a collaborative task. A number of endangered language projects bring together language activists, linguists, and computer scientists to work on dictionary creation. Web tools, such as the wiki WeSay,[12] allow native speakers to generate their own dictionaries with little technological fuss. For some communities, an online tool is problematic because it allows for seemingly uncontrolled public access to language resources and is subject to possibly unwarranted editing (Rau et al. 2009).

References

Abbi, Anvita. 2001. *A Manual of Linguistic Field Work and Indian Language Structures.* (Lincom Handbooks in Linguistics 17.) Munich: Lincom Europa.

Abney, Steve, Dunstan Brown, Osten Dahl,Sebastian Drude, Susanna Imrie, Marc Kemps-Snijder, Christopher Manning, Mike Maxwell, and Vivian Ngai. 2006. *Tools and Standards: The State of the Art Working Group 3: Lexicon creation*. E-MELD Workshop 2006 on Digital Language Documentation. Online: http://emeld.org/workshop/2006/wg/wg3-report.pdf

Abraham, Major R. C. 1964. *Somali-English Dictionary*. London: University of London Press.

Badten, Linda Womkon, Vera Oovi Kaneshiro, Marie Oovi, Christopher Koonooka, and Steven A. Jacobson. 2008. *St. Lawrence Island/Siberian Yupik Eskimo Dictionary*. Fairbanks: Alaska Native Language Center, University of Alaska.

Bartholomew, Doris A., and Louise C. Schoenhals. 1983. *Bilingual Dictionaries for Indigenous Languages*. Hidalgo, Mexico: Summer Institute of Linguistics.

Bee, Darlene and Alan R. Pence. 1962. Toward standardization of a survey word list for Papua and New Guinea. In *Studies in New Guinea Linguistics by Members of the Summer Institute of Linguistics* (New Guinea Branch) ed. by James C. Dean, 64-75. Oceania Linguistic Monographs, 6. Sydney: University of Sydney.

Beekman, John. 1975. Eliciting Vocabulary, Meaning, and Collocation. In *Language Learner's Field Guide*, ed. by Alan Healey, 361-86. Ukarumpa, Papua New Guinea: Summer Institute of Linguistics.

Béjoint, Henri. 2000. *Modern Lexicography: An introduction*. Oxford: Oxford University Press.

Berlin, Brent. 1992. *Ethnobiological Classification: Principles of Categorization of Plants and Animals in Traditional Societies*. Princeton, NJ: Princeton University Press.

Bouquiaux, Luc, Jacqueline M. C. Thomas, and James Robert. 1992. *Studying and Describing Unwritten Languages.* Dallas, TX: SIL International.

Breedlove, Dennis E., and Robert M. Laughlin. 1993. *The Flowering of Man. A Tzotzil Botany of Zinacantán*. Two vols. (Smithsonian Contributions to Anthropology No. 35.) Washington D.C.: Smithsonian Institution Press.

[12] http://wesay.org/wiki/Main_Page

Bridges, Thomas. 1933. *Yamana-English. A Dictionary of the Speech of Tierra del Fuego*, ed. by Ferdinand Hestermann and Martin Gusinde. Mödling, Austria: Missionsdruckerei St. Gabriel.

Canagaraj, A. Suresh. 1999. *Resisting Linguistic Imperialism in English Teaching*. Oxford University Press, Oxford.

Capell, Arthur 1965. *Linguistic Materials for Fieldworkers in Australia*. Canberra: Australian Institute of Aboriginal Studies.

Chafe, Wallace L., ed. 1980. *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. Norwood, N.J.: Ablex.

Coelho, Gail. 2006. Documenting plant and animal names. *Language Archives Newsletter* 8:2-6. Online: http://www.mpi.nl/LAN/.

Conklin, Harold C. 1962. Lexicographic treatment of folk taxonomy. *International Journal of American Linguistics* 28(2), Part IV. Problems in lexicography, ed. by Fred W. Householder and Sol Saporta, 119-41. Bloomington, IN: Indiana University Press.

Corris, Miriam, Christopher Manning, Susan Poetsch, and Jane Simpson. 2004. How useful and usable are dictionaries for speakers of Australian Indigenous Languages? *International Journal of Lexicography* 17:33-68.

Coward, David F., and Charles E. Grimes. 1995. *Making Dictionaries: a Guide to Lexicography and the Multi Dictionary Formatter, version 1.0*. Waxhaw, NC: Summer Institute of Linguistics.

Crowley, Terry. 2007. *Field Linguistics. A Beginner's Guide*. Oxford: Oxford University Press.

Cruse, D. Alan. 1986. *Lexical Semantics* (Cambridge textbooks in Linguistics). Cambridge: Cambridge University Press.

Day, Gordon M. 1994. *Western Abenaki Dictionary. Volume 1: Abenaki-English*. (Mercury Series Canadian Ethnology Service paper 128.) Hull: Quebec: Canadian Museum of Civilization.

Day, Gordon M. 1995. *Western Abenaki Dictionary. Volume 2: English-Abenaki*. (Mercury Series Canadian Ethnology Service paper 129.) Hull: Quebec: Canadian Museum of Civilization.

Dixon, R. M. W. 2007. Field Linguistics: a minor manual. *Sprachtypologie und Universalienforschung*. (*Focus on Linguistic Fieldwork*, ed. by Alexandra Y. Aikhenvald) 60(1):12-31.

Dixon, R. M. W. 2010. *Basic Linguistic Theory. Volume 1 Methodology*. Oxford: Oxford University Press.

Dixon, R. M. W., and Barry J. Blake, eds. 1979, 1981, 1983, 1991, 2000. *Handbook of Australian Languages*. 5 vols. Canberra: Australian National University Press (Vols. 1-3); Oxford: Oxford University Press (Vols. 4, 5).

Enrico, John. 2005. *Haida Dictionary*. (2 Vols.) Fairbanks, AK: Alaska Native Language Center.

Fleck, David W. 2007. Field linguistics meets biology: how to obtain scientific designations for plant and animal names. In *Sprachtypologie und Universalienforschung* (Focus on Linguistic Fieldwork, ed. by Alexandra Aikhenvald), 60(1):81-91.

Evans, Nicholas, and Alan Dench. 2006. Introduction: Catching language. In *Catching Language: The Standing Challenge of Grammar Writing*. (Trends in Linguistics. Studies and Monographs 167.), ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 1-39. Berlin: Mouton de Gruyter.

Foley, William A. 2002. Field methods. In *The Linguistics Encyclopedia*, Second edition, ed. by Kirsten Malmkjær, 131-137. London and New York: Routledge.

Franchetto, Bruna. 2006. Ethnography in language documentation. In *Essentials of Language Documentation*. (Trends in Linguistics. Studies and Monographs 178.), ed. by Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel, 183-211. Berlin and New York: Mouton de Gruyter.

Frawley, William, Kenneth Hill, and Pamela Munro, eds. 2002. *Making Dictionaries: Preserving Indigenous Languages of the Americas*. Berkeley, CA: University of California Press.

Gil, David and Yeshayahu Shen. 2004. Metaphors Across Languages. A Questionnaire. Online: http://lingweb.eva.mpg.de/jakarta/docs/English.pdf

Green, Jennifer. 2001. 'Both sides of the bitumen': Ken Hale remembering 1959. In *Forty Years on: Ken Hale and Australian Languages*, ed. by Jane Simpson, David Nash, Mary Laughren, Peter Austin and Barry Alpher, 29-43. Canberra: Pacific Linguistics.

Grenoble, Lenore and Lindsay Whaley. 2006. Orthography. Chapter 6 in *Saving Languages. An Introduction to Language Revitalization*, 137-159. Cambridge: Cambridge University Press.

Grimes, Joseph E. 1995. *Language Survey Reference Guide.* Dallas, TX: Summer Institute of Linguistics.

Guillaume, Gabriel. 1963-1966. *Atlas Linguistique Armoricain Roman*. Angers: the author.

Hartmann, R. K. K. 2007. *Bibliography of Lexicography*. Online: http://euralex.pbworks.com/f/Hartmann+Bibliography+of+Lexicography.pdf, 1962.

Haas, Mary R. 1962. What belongs in a Bilingual Dictionary? *International Journal of American Linguistics* 28(2), Part IV. Problems in lexicography, ed. by Fred W. Householder and Sol Saporta, 45-50. Bloomington, IN: Indiana University Press.

Hausmann, Franz Joseph, Herbert Ernst Weigand, and Ladislav Zgusta, eds. 1990-1991. *Dictionaries*. Vols 1-3. Berlin and New York: Mouton de Gruyter.

Haviland, John B. 2006. Documenting lexical knowledge. In *Essentials of Language Documentation*. (Trends in Linguistics. Studies and Monographs 178.) ed. by Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel, 129-162. Berlin and New York: Mouton de Gruyter.

Healey, Alan. 1964. *Handling Unsophisticated Linguistic Informants.* (Linguistic Circle of Canberra Publications, Series A, Occasional Papers No. 2.) Canberra: Linguistics, Department of Anthropology and Sociology, The Australian National University.

Healey, Alan, ed. 1975. *Language Learner's Field Guide*. Ukarumpa, Papua New Guinea: Summer Institute of Linguistics.

Heather, M. A., and B. A. Rossiter. 1988. Specialist dictionaries in electronic form. *Literary and Linguistic Computing* 3(2):109-121.

Henderson, Julius, and John P. Harrington. 1914. *Ethnozoology of the Tewa Indians*. Washington, D.C.: Government Printing Office.

Henne, Marilyn. 1991. Orthographies, language planning, and politics: Reflections of an SIL literacy muse. *Notes on literacy* 65.1. 1-18.

Hinton, Leanne. 2001. New writing systems. In the *The Green Book of language revitalization in practice*, eds. by Leanne Hinton and Ken Hale . San Diego: Academic Press. 239-250.

Holton, Gary. 2009. Relatively ethical: a comparison of linguistic research paradigms in Alaska and Indonesia. *Language Documentation and Conservation* 3(2):161-175. Online: http://nflrc.hawaii.edu/ldc/http://hdl.handle.net/10125/4424.

Hopi Dictionary Project, Bureau of Applied Anthropology, University of Arizona. 1998. *Hopi Dictionary/Hopiikwa Lavaytutuveni: A Hopi Dictionary of the Third Mesa Dialect with an English-Hopi Finder List and a Sketch of Hopi Grammar*. Kenneth C. Hill, Project Director and Editor-in-Chief; Emory Sekaquaptewa, Cultural editor; Mary E, Black, Associate Editor; Ekkehart Malotki, Senior Contributing Editor; Michael Lomatuway'ma, Contributing Editor. Tucson, AZ: University of Arizona Press.

Hyman, Larry M. 2001. Fieldwork as a state of mind. In *Linguistic Fieldwork* ed. by Paul Newman and Martha Ratliff, 15-33. Cambridge: Cambridge University Press.

Jäschke, H. A. 1881. *A Tibetan-English Dictionary. With Special Reference to the Prevailing Dialects. To Which Is Added an English-Tibetan Vocabulary*. London and Henley: Routledge & Kegan Paul.

Jetté, Jules, and Eliza Jones. 2000. *Koyukon Athabaskan Dictionary*. Fairbanks, AK: Alaska Native Language Center, University of Alaska.

Kari, James. 2007. *Dena'ina Topical Dictionary*. Fairbanks, AK: Alaska Native Language Center.

Karttunen, Frances. 1983. *An Analytical Dictionary of Nahuatl*. Austin, TX: University of Texas Press.

Kibrik, Aleksandr E. 1977. *The Methodology of Field Investigations in Linguistics. (Setting up the Problem.)* (Janua Linguarum, Series Minor, 142.) The Hague: Mouton.

Ladefoged, Peter. 2003. *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*. Malden, MA: Blackwell Publishing.

Landau, Sidney I. 2001. *Dictionaries: the Art and Craft of Lexicography*. Cambridge: Cambridge University Press.

Laughlin, Robert M. 1975. *The Great Tzotzil Dictionary of San Lorenzo Zinacantán*. (Smithsonian Contributions to Anthropology No. 19.) Washington D.C.: Smithsonian Institution Press.

Laughlin, Robert M., and John B. Haviland. 1988. *The Great Tzotzil Dictionary of Santo Domingo Zinacantán, with Grammatical Analysis and Historical Commentary*. Three vols. (Smithsonian Contributions to Anthropology No. 31.) Washington D.C.: Smithsonian Institution Press.

Levin, Beth C. 1993. *English Verb Classes and Alternations*. Chicago, IL: University of Chicago Press.

Man, Edward H. 1923. *A Dictionary of the South Andaman (Âkà-Bêa) Language*. Bombay: British India Press.

Masthay, Carl, ed. 1991. *Schmick's Mahican Dictionary*, with a Mahican historical phonology by David H. Pentland. (Memoirs of the American Philosophical Society, vol. 197.) Philadelphia: American Philosophical Society.

Masthay, Carl, ed. 2002. *Kaskaskia Illinois-to-French Dictionary*. Saint Louis, MO: Privately Published.

Mathiot, Madeleine. 1967. The place of the dictionary in linguistic description. *Language* 43(3):703-724.

Matisoff, James. 1978. *Variational Semantics in Tibeto-Burman: the 'Organic' Approach to Linguistic Comparison*. Philadelphia: Institute for the Study of Human Issues.

Matisoff, James A. 1988. *The Dictionary of Lahu* (University of California Publications in Linguistics, 111). Berkeley, CA and Los Angeles, CA: University of California Press.

Matisoff, James. 2003. *Handbook of Proto-Tibeto-Burman. System and Philosophy of Sino-Tibetan Reconstruction*. Berkeley, CA: University of California Publications in Linguistics, 135.

Matteson, Esther. 1976. Developing the lexicon of an unwritten language. *Notes on Translation* 59:18-27.

Maxwell, Michael, and William J. Poser. 2004. Morphological Interfaces to Dictionaries. In *Proceedings of COLING, 20th International Conference on Computational Linguistics. Workshop on Enhancing and Using Electronic Dictionaries*, ed. by Michael Zock, 65-68. Geneva: Coling. Online: http://acl/ldc.upenn.edu/coling2004/W10/pdf/13.pdf

Mayer, Mercer. 1969. *Frog, Where Are You?* New York: Dial Books for Young Readers.

Moore, Robert E. 1988. Lexicalization versus Lexical Loss in Wasco-Wishram Language Obsolescence. *International Journal of American Linguistics* 54(4):453-68.

Mosel, Ulrike. 2004. Dictionary making in endangered language communities. In *Language Documentation and Description*, vol. 2, ed. by Peter K. Austin, 39-54. London: School of Oriental and African Studies Working papers. Online: www.mpi.nl/lrec/2002/papers/lrec-pap-07-Dictionary_Endangered_SpComm.pdf

Mosel, Ulrike. 2006a. Grammaticography: The art and craft of writing grammars. In *Catching Language: The Standing Challenge of Grammar Writing*. ed by Felix Ameka, Alan Dench, and Nicholas Evans, eds., 11-68. Berlin: Mouton de Gruyter.

Mosel, Ulrike. 2006b. Fieldwork and community language work. In *Essentials of Language Documentation*. (Trends in Linguistics. Studies and Monographs 178.) ed. by Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel, 65-85. Berlin and New York: Mouton de Gruyter.

Nichols, Johanna, and Ronald L. Sprouse. 2003. Documenting lexicons: Chechen and Ingush. In *Language Documentation and Description*. Volume 1. ed. by Peter K. Austin, 99-121 London: SOAS, Hans Rausing Endangered Languages Project.

Nida, Eugene A. 1958. Analysis of meaning and dictionary making. *International Journal of American Linguistics* 24:279-292.

Parks, Douglas R., and Lula Nora Pratt. 2008. *A Dictionary of Skiri Pawnee*. Lincoln and London: University of Nebraska Press.

Pittman, Dean. 1948. *Practical Linguistics. A Textbook and Field Manual of Missionary Linguistics*. Cleveland, Ohio: Mid-Missions.

Poser, William J. 2002a. Lexical Databases for Carrier. Unpublished manuscript, University of Pennsylvania. Online: http://billposer.org/Papers/carlexdb.pdf

Poser, William J. 2002b. Making Athabaskan Dictionaries Usable. (Proceedings of the 2002 Athabaskan Languages Conference) Alaska *Native Language Center Working Papers* 2:136-147. Fairbanks, AK: Alaska Native Language Center. Online: http://billposer.org/Papers/makath.pdf

Rau, Victoria, Meng-Chien Yang, Hui-Huan Chang, and Maa-Neu Dong. 2009. Online Dictionary and Ontology Building for Austronesian languages in Taiwan. *Language Documentation and Conservation* 3(2):192-212. Online: http://nflrc.hawaii.edu/ldc/http://hdl.handle.net/10125/4439.

Redard, Georges. 1960. *Atlas linguistique de l'Iran: Questionnaire normal*. Berne, published by the author.

Roberts, James, and Keith Snider. 2005-2006. SIL Comparative African Wordlist (SILCAWL) *SIL Electronic Working Papers 2006*. Online: http://wwwsilewp

Rudes, Blair A. 2002. Early Vocabulary and Dictionary Development. A Cautionary Note. In Indigenous Languages Across the Community ed. by Barbara Bunaby and Jon Reyhner, 189-195. Flagstaff, AZ: Northern Arizona University. Online: http://jan.ucc.nau.edu/~jar/ILAC/ILAC_20.pdf

Sapir, Edward. 1931. *The Southern Paiute Language: Southern Paiute Dictionary*. Proceedings of the American Academy of Arts and Sciences 65(3):537-730.

Sapir, Edward, and Morris Swadesh. 1960. *Yana Dictionary*, ed. by Mary R. Haas. (University of California Publications in Linguistics 22.) Berkeley, CA and Los Angeles, CA: University of California Press.

Seifart, Frank. 2006. Orthography development. In *Essentials of Language Documentation*, ed. by Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel, 275-299. Berlin: Mouton de Gruyter.

Smedt, A., de, C. I. C. M., and Antoine Mostaert, C.I.C.M. 1933. *Le Dialecte Monguor Parlé par les Mongols du Kansu Occidental*. IIIe Partie: Dictionnaire Monguor-Français. Pei-p'ing: The Catholic University of Peking.

Stanford Linguistics Fieldwork Committee. 2006. Online: http://www.stanford.edu/dept/linguistics/cgi-bin/fieldwork/contactus.php

Stebbins, Tonya. 2003. *Fighting language endangerment: community directed research on Sm'algyax (Coast Tsimshian)*. (Endangered Languages of the Pacific Rim A2-026.) Osaka: Faculty of Informatics, Osaka Gakuin University.

Sterkenburg, Piet van, ed. 2003. *A Practical Guide to Lexicography*. (Terminology and Lexicography Research and Practice 6). Amsterdam and Philadelphia: John Benjamins.

Sutton, Peter, and Michael Walsh. 1979. *Revised Linguistic Fieldwork Manual for Australia*. Canberra: Australian Institute of Aboriginal Studies.

Sutton, Peter, and Michael Walsh. 1987. *Wordlist for Australian Languages*. 2nd edition. Canberra: Australian Institute of Aboriginal Studies.

Svensén, Bo. 2009. *A Handbook of Lexicography. The Theory and Practice of Dictionary-Making*. Cambridge: Cambridge University Press.

Sylestine, Cora, Heather K. Hardy, and Timothy Montler. 1993. *Dictionary of the Alabama language*. Austin, TX: University of Texas Press.

Swadesh, Morris. 1955. Towards greater accuracy in lexicostatistic dating. *International Journal of American Linguistics* 21:121-137.

Thompson, Nile. 1989. On the Need for Ethnographic Information in Native American Dictionary Construction. *Papers of the 24th International Conference on Salish and Neighboring Languages*, ed. by Nile Thompson, 67-82. Steilacoom, WA.

Ullrich, Jan, compiler, editor and author. 2008. *New Lakota Dictionary. Lakȟotiyapi-English/English-Lakȟotiyapi and Incorporating the Dakota Dialects of Yankton-Yanktonai & Santee-Sisseton*. Bloomington, IN: Lakota language Consortium.

Vaux, Bert, and Justin Cooper. 1999. *Introduction to Linguistic Field Methods.* (Lincom Course books in Linguistics 1). Munich: Lincom Europa.

Vaux, Bert, Justin Cooper, and Emily Tucker. 2007. *Linguistic Field Methods*. Eugene, OR: Wipf & Stock.

Wolff, Hans. 1969. *A Comparative Vocabulary of Abuan Dialects*. Evanston, IL: Northwestern University Press.

Zgusta, Ladislav. 1971. *Manual of Lexicography*. The Hague: Mouton.

Chapter 10
Phonetic and Phonological Fieldwork

Fieldwork is like heart surgery: you can learn to do it well only by practicing on someone. (Ladefoged 2003, Acknowledgements)

Peter Ladefoged has said that there is, "nothing more ephemeral than the sounds of a language. The sounds will live only as long as the language is spoken. When the sounds are those of elderly speakers whose children belong to another world, then soon those sounds will be gone forever. All that can remain are whatever records we have been able to archive."[1] Understood in this light, we recognize that efficient recording and accurate phonetic transcription is of the utmost importance for language documentation. Also, we know that a carefully transcribed corpus is the basis for reliable phonological analysis. So whether the goals are language documentation or language description, there is no room for sloppy phonetic fieldwork.

10.1 Preparing for Phonetic Fieldwork

A first step in discovering the phonetic inventory of a language is to collect any previous descriptions or documents on the language. These may be phonetic and phonological descriptions, or even texts, since the goal is to become acquainted with what can be expected in terms of sounds, syllable structure, phonotactics, and suprasegmental features. It would also be useful to listen to any available recordings of the language. Earlier transcriptions of recordings should be used with care, since they might be inconsistent, describing a dialect different than the one under consideration, or they could represent a combination of dialects or languages (Ladefoged 2003:2).[2] It is also recommended that the fieldworker contact other linguists who have worked on the target language for information on useful sources and materials of phonetic interest (Gordon 2003). Finally, the beginning

[1] http://www.linguistics.ucla.edu/people/ladefoge/Preserving%20sounds.pdf

[2] It may be necessary to apply for special permission to work with previous recordings. In addition, as discussed in Bowern (2008:186–187), there are specific data management challenges related to working with early recordings.

S.L. Chelliah and W.J. de Reuse, *Handbook of Descriptive Linguistic Fieldwork*,
DOI 10.1007/978-90-481-9026-3_10,

fieldworker might review phonetic and phonological concepts from a discussion such as Dixon (2010a:264–288).

10.2 Organizing Word Lists and Short Phrases

The fieldworker will want to prepare word lists and short phrases to help with phonetic investigation. Since these lists are used to determine the sounds of the target language, they should be recorded in narrow phonetic transcription. The reader should refer to Section 9.2.1 for a discussion of word list development. With a list of 500–700 words, it will generally be possible to develop an inventory of the phones of the target language. For languages with smaller phonetic inventories or where the phonetic inventory of related languages is known, a shorter word list may suffice.

Why not just record a large quantity of naturally occurring speech and determine the sounds from those data? This might be possible if the sound system of closely related languages has already been described. However, it is hard to recover such information from naturally occurring speech because it takes an inordinate amount of work to isolate similar phonetic environments for the comparison of sounds. Also, transcription of natural data is difficult, and it is useful to have some idea of the phonology before beginning this task (Gleason 1961:287).

The challenges of using word lists in phonetic work are: (1) to get the most natural pronunciation possible for acoustic analysis; (2) to record a number of words so that they are conveniently organized for instrumental analysis; and (3) to create data sets that position target sounds in a variety of phonetic environments. We will consider each of these in turn.

List intonation can obscure natural pronunciation. The last word in a list typically has lower pitch than earlier words in the list. The vowels of the list-final words are also often lengthened. On the other hand, words that occur earlier in the list may take on a uniform rising pitch at the end of the word. To get a pronunciation without these list intonation effects, it is recommended that list structure – beginning, middle, and end – be manipulated. For example, for recording purposes, a consultant might be asked to repeat a list in which words that occur at the beginning of the list are repeated at the end (Ladefoged 2003:7).

In addition to recording words in isolation, words should be recorded in a frame or carrier sentence. This is because articulation is affected by adjacent segments through assimilation. These effects should be isolatable across the word sets being compared. So, for example, when studying vowel length, vowels should be recorded next to voiced consonants and voiceless consonants so that the effects of voiced consonants on vowel length can be taken into consideration. Additionally, frames will help circumvent the list intonation effects discussed above.

A well-constructed frame will allow for easy extraction and comparison of sounds across tokens of the same or different words. For example, Ladefoged

(2003:7) suggests that, when studying consonant length, the target consonant should be sandwiched between two vowels. This way there is a clear beginning and end point for the consonant. A typical frame for this in English would be "Say _____ again". Different frames should be used depending what the object of study is. So, when studying vowel length, the target vowel being studied should be sandwiched between two consonants, thereby providing a clear beginning and end point for the vowel. A typical frame for this in English would be "Repeat ____ twice".

The frame itself should not contribute to sandhi effects. For a verb final language, a common frame is "I say the word ___"; the target word will then be in utterance-medial position, since the verb 'say' will occur in final position. Ladefoged (2003:7) also cautions that the frame used should place the target word in a position where it receives nuclear stress.[3] This will ensure a consistent environment in terms of vowel length, stress, and pitch for each word studied in the minimal sets. When creating frames it may also be useful to compare features in words produced in isolation and within frames. For example, Hildebrandt (2005:27–28) shows that in words recorded in isolation the vowels undergo lengthening when in unbounded contexts, i.e. open syllables or not followed by other segments, but the same words recorded in frames do not show such lengthening.

10.3 Recording Word Lists and Short Phrases

To make a clean and clear recording for further instrumental analysis the field-worker must minimize the transfer of environmental noise to the recording. This can be done in two ways: (1) controlling the recording environment; and (2) using equipment that does not pick up background sound.

One way of controlling the recording environment is to turn off machines that hum or buzz. In India, even in a room with all the windows closed, one can usually hear the sound of the ceiling fan or the hum of an air conditioner. I [Chelliah] often shut these off just while recording for phonetic fieldwork, then turn them back on for the rest of the session. This does not help in cancelling the noise of the vegetable vendor crying out what he has for sale that day! Since we filter out the background noise that forms part of our daily lives, the sounds of the vegetable vendor, bird song, or the call for morning prayers are only noticed when we are recording. Austin (2006:90) lists four ways to manage background interference: (1) use headphones when recording to monitor what is being recorded; (2) use a unidirectional microphone that will only pick up sound directly in front of the microphone; (3) use a microphone shield in windy conditions, and move the microphone as close to the speaker's mouth as possible while still retaining good recording quality;

[3] Also called 'sentence stress', this refers to the main stress appearing within a single phrase.

and (4) restrict unnecessary movement of people and materials. Ladefoged (2003:21) recommends that recordings should not be made in rooms that are sparsely furnished, since in empty rooms sound waves will reverberate off walls. Alec Coupe (p.c.) reports recording in an open field on a windless day to reduce the chance of echo. There is a trade off: recording the sounds of wood-choppers, croaking frogs, or cicadas.

If possible, a soundproof booth at a local university or recording studio could be used.

10.4 Finding Appropriate Consultants for Phonetic Fieldwork

The physical condition and age of the native speaker must be considered when selecting consultants for phonetic fieldwork. Ideally, the native speaker should have all their own teeth, be free of speech impediments such as a stammer or lisp, and not have hearing problems. Older speakers have less control over their articulators (Maddieson 2001:218), but this doesn't mean they be used for phonetic fieldwork since, as Abbi (2001:77) notes, older consultants may be the only speakers to have retained a particular word with an unusual sound. Also, if the speech community is sensitive to the idea of "correct" pronunciation and feel that only certain elders possess that pronunciation, then working with elders is crucial. Ladefoged (2003:22–23) suggests having a good speaker say the word first and then getting others to repeat or comment on that word to get a community-approved pronunciation of the word. Therefore, both young and old speakers should be recorded. Recordings from younger speakers will be the most useful for instrumental study, while participation of older speakers will endorse the project as worthwhile, provide variant pronunciations, and support younger speakers in their participation (Maddieson 2001:218). Finally, Yvonne Treis (p.c.) reports that blind informants make good consultants for phonological fieldwork because their sense of phonetic distinctions is heightened.

How many speakers should data be collected from? For phonological analysis, the fieldworker should begin with a single speaker, and then compare results with data from other speakers. The fieldworker should keep track of who has been recorded, and collect basic biographical information from each. For phonetic analysis of sounds, the analysis should always be based on several tokens and several speakers; Ladefoged (2003:14) recommends from 24 to 40 speakers, half of these male and half female. Gordon (2003) recommends six male and six female speakers as an ideal pool to collect data from for quantitative study. An interesting point made by Hildebrandt (2007) is that in politically unstable situations, such as in Nepal, where a guerilla insurgency has depleted the younger population from villages, selecting the recommended number of speakers and finding younger speakers may not be possible.

10.5 Requirements of the Native Speaker

The native speaker should obviously not be expected to list the sounds of her or his language. Native speakers can tell you what is NOT a sound in their language, but it is hard for speakers to provide a comprehensive list of sounds. The native speaker can perceive phonetic distinctions when they are pointed out, but phonetic and phonological information is not readily accessible at a conscious level. Most speakers are surprised when phonemes or their allophonic variants are pointed out. Bilingual speakers may be able to report on sounds they consider unusual in the target language.

In early sessions, the fieldworker will not be not accustomed to hearing the target language. Additionally, s/he will be preoccupied with identifying and carefully transcribing the sounds of the language, so s/he will certainly need to ask for utterances to be repeated. Healey (1964:350) notes that excessive requests for repetition will fatigue or bore the native speaker. When a speaker is asked to repeat a word several times, there's a point where s/he will start to "disassociate" the meaning from the sound. When that happens, the contrasts being produced are unreliable. Repetition fatigue and boredom will blur intonation patterns and allophonic distinctions. Repetition fatigue is also a problem for the fieldworker because the ear slowly dulls to phonetic contrasts. Finally, a request for repetition may be interpreted by the speaker as a sign that s/he is doing something wrong and is being asked to correct that error in the second telling.

In order to prevent repetition fatigue, the fieldworker might limit requests for repetition to three or four tokens. If a particular phoneme or allophone is problematic, this can be checked over a period of time rather than cramming questions into a single session. An advantage of the "taking it slow" approach to pinning down problematic sounds is that the fieldworker has more opportunity to hear them in various contexts.

To reduce consultant fatigue and frustration, the fieldworker should try to pronounce words that include the target sound, and the consultant should be trained to help identify exactly what part of the fieldworker's pronunciation is wrong. Finally, the fieldworker might listen to or instrumentally analyze the recorded word rather than asking the native speaker for more repetitions. This is an especially easy tool for studying voicing contrasts (as measured by voice onset time, or VOT), as well as vowel quality, and pitch.

It is also important to understand the nature of repetition. When a speaker repeats a word, s/he may indeed intend to give the fieldworker a second chance to catch the sounds. Sometimes, however, the speaker might actually be providing two different forms in order to contrast a particular sound. The fieldworker may not recognize whether the speaker is producing a correction, a different utterance, or a repetition.

10.6 Transcription

High quality transcription is of vital importance, so it is a good idea to do some practice transcription before heading for the field. An easy way of doing this is to use the resources on the UCLA Phonetics Lab Archive (2007) or the web pages which accompany Ladefoged (2005a, b). Here, one can listen to sounds from a variety of languages and learn the correlation between IPA symbols and the sounds they represent. It is useful to practice transcribing the contact language and reading about the language family of the target language to see what sounds are usually found in these languages and how they are transcribed or written in the standard orthography. This more general “ear training” (Redden 1982) should be followed with language specific ear training in the field. Healey (1964) has some suggestions for how this can be done. Fieldworkers should surround themselves with natural input and find situations in which the target language is used so as to become accustomed to hearing it, and if at all possible become proficient in some conversational routines (e.g. Greetings, buying food). It is useful to record words and short sentences and convert these to phonetic drills that can be practiced through repetition. These repetitions can then be recorded and the fieldworker’s pronunciations can be compared with native pronunciations.

A common experience reported by fieldworkers is a tentativeness or skepticism about their transcriptions in the early stages of fieldwork. It is best to trust one’s ear and instincts about what is heard and how it is transcribed. Many times these early transcriptions turn out to be accurate, and later “corrected transcriptions” turn out to be wrong! There will always be a set of sounds which the fieldworker will have persistent doubt about and which will need further analysis.

As work progresses, the fieldworker will begin to hear and transcribe with greater speed. At this point, in order to be consistent in transcriptions, there is a temptation to move quickly to phonemic transcription, thereby collapsing possibly significant distinctions. So, when should a fieldworker switch from narrow to broad transcription? It really depends. The fieldworker should only move quickly to a phonemic transcription system or adopt a practical orthography early on in fieldwork if the basic phonetics and phonology of the language have already been established by previous fieldwork. One reasonable way to proceed is to continue using narrow transcription only for those sounds that are problematic, i.e. that vary in phonetic detail. A good overview of the different levels of transcription is given in Laver (1994:549–562). Of particular interest are the examples of narrow phonetic transcription of words and phrases in a number of languages (558–560).

There are compelling reasons for phonemicizing as soon as possible. First, by using phonemic transcription, the fieldworker commits to a hypothesis which can then be tested. Bowern (2008:64) lists these advantages of phonemic transcription: (1) it is more efficient than phonetic transcription, because in a phonemicized system there are fewer diacritics to write or type, and a smaller variety of them; (2) phonemic transcription facilitates computerized searches: if transcription is phonetic, it is that much harder to look for all occurrences of a morpheme; and (3) phonemic spelling is also preferable for the publication of community texts. As discussed by Sutton and Walsh (1979:16), it is useful to continue with narrow transcription if speakers consistently vary in their pronunciation of certain words. These transcriptions might eventually help establish that the consultants speak different dialects.

10.7 Phonetic and Phonological Analysis

One factor to be aware of in phonetic and phonological analysis is that the fieldworker's perceptions are naturally influenced by the sounds and sound patterns of his/her native language, or other languages in which s/he is dominant. This interference can be particularly acute when the fieldworker's native or dominant language is very similar to the target language.[4] Careful transcription and construction of word lists and phrase lists can help counteract such interference effects. Here are some useful examples, taken mainly from Bowern (2008:41), of sounds that native speakers of English have difficulty in perceiving:

- Aspiration in codas, and aspiration in fricatives or affricates is missed.
- Bilabial and labio-dental fricatives are not distinguished.
- Diphthongs and disyllabic sequences of vowels are not distinguished.
- The distinction between syllable initial voiceless unaspirated stops and voiced stops is missed.
- Double articulations such as [kp] are perceived as a single articulation, i.e., [k]. See Gleason (1961:282) for a list of double articulations that are mistakenly transcribed as segment sequences.
- Glottal stops are hard to perceive or are ignored.
- Glottal, uvular, or velar fricatives are not distinguished.
- Nasalization is not perceived, or is perceived as a vowel–nasal sequence.
- Palatal lateral or nasals are perceived as sequences like [lj], [nj].
- Palatal stops and palatal affricates are not distinguished.
- Retroflex, alveolar, and dental points-of-articulation are not distinguished.
- Syllable initial velar nasals can be perceived as [m], [n], or [w].
- Uvular consonants are not distinguished from velar consonants.
- Voiceless laterals and trills are perceived as [s], [ʃ] [sl], [ʃl], or [θl].
- Voiceless vowels are heard as /h/, perceived as an extra puff of air, or are missed completely.
- Vowel length is ignored or transcribed as stress or tenseness.
- Vowels [ɪ] and [ɛ] merge before nasals in some English dialects; thus distinctions between these vowels will be hard to hear.

To this list we could add tonal contrasts.

Even with training, fieldworkers may come across a sound that they can perceive but not transcribe. If this happens the fieldworker could choose a diacritic that best describes the situation at hand, e.g. a raised *h* for murmur or heavy aspiration in unexpected places such as after fricatives. Sutton and Walsh (1979) suggest that if one is totally unsure about what it is that makes a sound seem unusual, it can be written

[4] See Krishnamurthy (2007) for a good example of how a linguist must distinguish his or her native language from a closely related language being studied. In this article, Krishnamurthy describes and distinguishes Konda, a Dravidian language, from his native language Telugu, also a Dravidian language.

down using the symbol that most closely represents the sound heard and then underlined. Phonetic details should not initially be ignored, since a pattern, if it exists, will only become clear when occurrences are noted. On the other hand, one should not go too wild in the use of diacritics since they can become hard to keep track of.

With a list of several hundred words, it should be possible to hypothesize which of the transcribed phones are most likely phonemes of the target language. If particular sounds have wide distribution and are known in related languages to be phonemes, then the fieldworker can initially assume that these sounds are phonemes in the target language. Hypotheses can also be based on general typological tendencies. For example, if two voiceless stops [p] and [t] are phonemes, then the third, [k], most likely is as well. Yet the fieldworker should remember that phonological patterning is not always very neat. For example, Dutch has the phonemes /p/, /t/, /k/, /b/, and /d/. But [g] is an allophone of /k/. In Arabic, there is a /b/ but no /p/; there is also a /f/ but no /v/. The fieldworker's working hypothesis will be revised as more data is added to the corpus. We are sometimes concerned about jumping to conclusions without clear evidence; however, as pointed out by Gleason (1961:296), hypothesis formation in fieldwork is an absolute necessity to move the investigation forward. There must be analysis and hypothesis about language structure at every step of fieldwork. Additionally, new phonemes may be noticed quite late in fieldwork – some report finding new phonemes 8 or 9 months into fieldwork (Gerd Jendraschek p.c.) – in which case revisions of hypotheses will take place as new data come to light.

10.7.1 Collecting Minimal Pairs If Possible

If the target language is monosyllabic or non-fusional (that is, bound morphemes can be easily separated from the root or stem), then minimal pairs or minimal sets can be easily collected from the word lists to illustrate sound contrasts. Often, in such cases, speakers with no prior training can list contrasting words.

Sounds in word-initial position will be the easiest to contrast because these are readily recalled by speakers (Maddieson 2001:217), but obviously contrasts may exist word medially and finally as well. Near minimal pairs can be used to establish contrasts when minimal pairs cannot be found. Rare sounds might be found only in archaic or borrowed words, and so a literate or older speaker may be required to produce or pronounce these words (Abbi 2001:77).

When the fieldworker is unsure whether or not two sounds contrast, s/he could try PAIR TESTING, a method discussed by Harris (1951). The method is as follows. Take two words that appear to contrast a sound. Repeat these to a speaker an equal number of times, embedded in a list of about 40 other minimal pairs. If the speaker consistently assigns the same two distinct meanings to the two words close to 100% of the time, then one can reliably conclude that the target sounds in the word pair are indeed distinct phonemes. If the same meaning is assigned to both words 50% of the time, the two sounds in the word pair are in free variation.

Meaning is important to setting up a minimal pair. It is not enough to put together a string of words and ask if they mean the same thing or not. The native speaker must be asked to assign a meaning to the words in question. An illustration of why a definition must be established is the following. In Manipuri, *thá* means 'drop', 'send', 'release', or 'give up.' However, some consultants may not be able to accept that a single form has such seemingly irreconcilable meanings as 'send' and 'give up'. When confronted with such polysemy, I [Chelliah] have noticed that some speakers impose stress or different tones on what is actually the same word, in order to bolster their "these are two different words" analysis. It is important for fieldworkers to have information that will allow them to make a decision on whether there are in fact two words – i.e. a real minimal pair, contrasting in stress or tone or the same word, produced by the consultant after folk phonological over-analysis.

Minimal pairs should not be the only way, or even the first way, to establish phonemes, since this can be time consuming. For acoustic analysis, it is useful to have sounds in lists where the phonetic environment is comparable. For example, if voiced stops are being studied, it will be necessary to see those stops adjacent to a variety of vowels and consonants (phonotactics allowing), and in word-initial, -medial and -final position. Minimal pairs can be culled from these lists.

Model structuralist phonemic analyses to consult are: Ferguson (1960) on Bengali; Matthews (1955) on a Dakota dialect (Siouan); and Haugen (1958) on Icelandic.

10.7.2 Comparing Like Sounds in Various Phonetic Environments

In order to identify sounds that the fieldworker suspects may be allophones of the same phoneme, the distribution of *like* consonants and vowels in a number of phonetic environments will have to be investigated. It is helpful to work from more general environments to more specific environments when doing this sort of analysis: so like consonants could be compared first adjacent to vowels, then adjacent to high vowels, and finally, adjacent to high front vowels.

Since two sounds may not contrast in all environments, it is important to check their distribution in several different environments. Gleason (1961:291) gives the following example: A fieldworker records oral stops in word initial position in English. She gets three voiceless aspirated stops [p^h, t^h, k^h] and three voiced stops [b, d, g]. She may be led to believe that these sounds are in contrastive distribution. However, when she checks word-medial position, she learns that [p, t, k] contrast with [b, d, g], with the voiceless stops only moderately aspirated. From this she learns that the word-medial environment is needed to get an accurate picture of the distribution of stops in English.

When phrases are used for phonetic analysis, it may be difficult to identify word boundaries. Vaux and Cooper (1999:26) suggest that stress can sometimes be used to identify the beginning of a word, since – cross-linguistically – primary stress often falls on the first or last syllable of a word. Native speaker input on phonological and grammatical word boundaries may provide some insights, but the fieldworker should not depend on this wholly since it is not always easy to determine what a word is (Dixon and Aikhenvald 2002; Dixon 2010b:1–36). Orthography can also influence word boundary judgments. Manipuri, for example, has frequent combinations of verbal suffixes that are sometimes written as separate words; therefore some speakers may represent these to the fieldworker as separate forms. Additionally, because words can get up to ten syllables long in Manipuri, a syllable with secondary stress within a long word is sometimes represented orthographically as the start of a new word. In some languages, such as Burmese, the verbal suffixes are written as separate words. So, the writing system helps in morphological analysis but not in determining word boundaries.

10.7.3 Formulating Hypothesis, Preparing Questions

On the basis of a list of sounds and the environments where these sounds occur, the fieldworker can formulate an initial hypothesis about the contrastive and complementary distribution of sounds in the target language. The fieldworker might want to refer to Gleason (1961:271–285) or Harris (1951:60–78), which provide methodologies for distributional analysis for fieldworkers. Common types of alternations to look for are:

- Cluster simplification
- Coronal Harmony
- Epenthesis
- Final devoicing of consonants
- Spirantization (fricativization) of stops between vowels
- Glide insertion between vowels
- Place assimilation (especially for nasals)
- Released and unreleased stops
- Voice assimilation (especially in consonant clusters, in intervocalic position, and following nasals)
- Vowel devoicing between voiceless consonants or word-finally
- Vowel harmony
- Vowel reduction (peripheral vowels become centralized when unstressed)
- Vowel rounding, backing, or fronting when adjacent to round, back, or front consonant articulations

The following is a useful list of common allophones from Gleason (1961:275):

- Corresponding voiced and voiceless sounds, e. g. [k g], [s z]
- Corresponding stops and fricatives, e. g. [k x]
- Bilabials, labiodentals, bilabials and labiodentals, stops, fricatives, nasals, or stops and fricatives, e. g. [β v], [m ɱ], [Φ f]
- Dental, alveolar, and retroflex stops, fricatives, and laterals, e. g. [t l]
- Stops and fricatives, e. g. [t, θ]

- Alveolars and alveopalatals, e. g. [s ʃ]
- Palatal and velar stops or fricatives, e. g. [c k]
- Corresponding aspirated and unaspirated stops, e. g. [tʰ t]
- All nasals(except that [m] is usually distinct), e. g. [n ŋ]
- All varieties of [r]-like sounds and many [l] like sounds
- Dental and alveolar flaps and stops, e.g. [d ɾ]
- Uvular [R] and velar or uvular fricatives, e.g. [ɣ ʁ]
- [h] and all unvoiced velar or palatal fricatives, e.g. [h x]
- All pairs of adjacent vowels, e.g. [ɪ ü] [ɨ i] [æ a] [o ɔ][5]
- Semivowels (glides) and labial, palatal or velar fricatives, e.g. [w ʋ]

Abbi (2001:96) and Bowern (2008:38) explain how an IPA chart can be of use in keeping track of ongoing analysis. For example, the fieldworker could:

- Circle the sounds on an IPA chart as they are discovered
- Draw boxes around sounds on an IPA chart that seem to be allophones
- Make note of odd gaps in articulatory features so those can be checked
- Make note of unexpected sounds, so those can be checked

10.7.4 Accessing Native Speaker Input

Hoijer (1958) suggests that to some extent the fieldworker must learn the target language so that his or her own intuitions can be used in determining if sounds are in contrastive or complementary distribution. But he also notes that native speaker input can considerably speed up phonemic analysis. The following is a discussion of how hypotheses about sound distribution can be confirmed by questioning the native speaker.

10.7.4.1 Checking Transcriptions

Obviously, native speaker input is necessary for checking transcriptions. As discussed in Ladefoged (2003:11), there are several ways to do this. First, words should be pronounced back to the consultant as they are elicited, and the consultant should be asked to correct the fieldworker's pronunciation as necessary. Then the fieldworker should read back words from transcriptions and see if the native speaker can understand what is being said. If the consultant says that the word is being mispronounced, the fieldworker should try to discern which part of the word they are mispronouncing:

[5] Note that it is often easier to identify phonemic vowel quality when there are more vowel phonemes than when there are only two or three vowels because in these cases a vowel may have a variety of phonetic instantiations. For example, /i/ may have the allophones [i], [ɪ], [e] and [ɛ].

is it the consonant or the vowel? is it the tone or the voice quality? To do this, the field worker could produce two variants, altering a specific feature, and ask which one is better. If neither is right, then perhaps some other feature, in another part of the word, is being mispronounced. Once the speaker accepts a pronunciation, the fieldworker can re-transcribe the word; however, s/he should not cross out the original transcription. At a later date the original and retranscribed forms can be checked with a consultant. Does the consultant still approve of the second pronunciation? Sometimes, it is possible to discover that both first and revised transcriptions are acceptable to speakers. Comparing alternate transcriptions of tokens can also help with the analysis.

The native speaker's articulators – for example lip rounding or spread lips – can provide useful guides to transcription. Speakers' suggestions on how to articulate sounds can also be instructive. Speakers often have their own terminology for talking about articulation and for identifying sounds. For example, English speakers often talk about "hard" and "soft" *c*. Rather than teach speakers linguistic terminology, it is generally easier to adopt their terminology. If the native speaker can't describe the position of their articulators, the fieldworker could offer some suggested descriptions to choose from (Bowern 2008:67). Advanced study of articulation is discussed in Ladefoged (2003:30–31) and Maddieson (2001:223). The possibilities include:

- Still and video photography: Used to determine the exact place of articulation for sounds produced at the lips and teeth.
- Palatography: Used to determine contact between the tongue and the palate. A description of this procedure, along with detailed instructions, is available in Ladefoged (2003:30–51) and at The UCLA Phonetics Lab website.[6] Vaux and Cooper (1999:66) discuss another procedure for studying articulation at the palate, where sequences of tongue touches to the palate are transmitted to a computer through electrodes connected to a mouth guard. Arka (n.d.) is a recent field study using palatography.
- Linguograms: Used to determine what part of the tongue is active in articulation. See Ladefoged (2003:40–42).
- Ultrasound: A portable ultrasound unit can be used to record tongue shape during articulation. The ultrasound images can be captured by a video recorder and used along with an audio recording for analysis. See E-MELD (2005) and Gick (2002) for more discussion.

Aerodynamic study recording air pressure and airflow can be used, among other things, to study features such as nasality versus orality; breathiness versus voicing; and airstream mechanisms used for consonant production. The required equipment and procedures are discussed in Ladefoged (2003:54–74). Some of these procedures can be done in the field. Others require local anesthesia and must be performed by a clinician.

[6] http:/www.humnet.ucla.edu/humnet/linguistics/faciliti/facilities/physiology/epg.html

Perception studies are also useful for obtaining native speaker judgments. For example, to determine to what extent a speaker's perception of consonant articulation is affected by surrounding vowels, the fieldworker can edit natural speech, creating sequences where different vowel qualities occur before and after the target consonant. Consultants' identification of consonants in these spliced sound files tell the fieldworker how vowel quality affects consonant perception. An example of the results of a perception study is as follows: when the [s] preceding a voiceless unaspirated stop – as in the word *skunk* – is instrumentally removed, native speakers of English hear a voiced stop, as in the word *gunk*.

Literate or linguistically sophisticated speakers who recognize the English alphabet or the IPA may well try to influence the fieldworker's transcription. It is best to write down transcription insights from a consultant. For one thing, it is important to show a consultant that their opinions are taken seriously. Secondly, native speaker comments can provide valuable information about the phonological system of the language and its relationship to orthography. We have already noted this with regard to word boundary judgments. When doing fieldwork on Manipuri, for example, one consultant insisted that I [Chelliah] transcribe as [ʃ] two sounds that I heard as [s^h] and [$ʃ^h$]. I learned later that the spelling system represented words pronounced with [s^h] or [$ʃ^h$] using the same grapheme, and that [s^h], [$ʃ^h$], and [ʃ] are in free variation for many speakers. We learn from such examples that native speaker comments should always be noted; don't buy into the notion that native speakers are naïve speakers. Some inconspicuous method should be used to keep a record of the fieldworker's transcription and the native speaker's input.

10.7.4.2 Checking Hypotheses

In addition to the two tests discussed in Section 9.2.2, Harris (1951:32–33) describes the DIFFERENTIAL STATUS TEST, which uses native speaker intuitions to determine the phonemic status of sounds. To perform this test, the researcher pronounces two variants of a word: variant 1 with sound A, and variant 2 with sound B. If it is suspected that sounds A and B are allophones of the same phoneme, the researcher would then ask the native speaker if the sounds in variants 1 and 2 are the same or different. The two possible results are:

- The speaker identifies the sounds as being the same. In this case, the sounds that have been transcribed differently are allophones of the same phoneme. Perhaps they are in free variation.
- The speaker identifies the sounds as being different. In this case, the sounds that have been transcribed differently are indeed allophones of different phonemes.

This is quite different from asking the native speaker directly what the status of a sound is. Vaux and Cooper (1999:64) list this "direct approach" as a last resort; as they point out, a native speaker might under-report distinctions if s/he is influenced by spelling, which often collapses distinctions. We know, for example, that bilingual speakers collapse distinctions in one language to fit with the phonemic inventory of another. Gudschinsky (1958:341) gives this example for speakers of Mazatec (Otomanguean of Mexico) and Spanish: Since Spanish does not distinguish between [t] and [tj], Mazatec speakers lump [t] and [tj] together when talking about the status of these sounds in Mazatec, too. So, one consultant judged [tkwa] 'hail' and [tjkwa] 'snail' as the same except for tone. In fact, the words have the same tone, but different initial consonants. The reverse is also possible: if two sounds are separate phonemes in the contact language, the consultant might regard them as separate phonemes in the target language as well, even if they are, in fact, allophones of the same phoneme.

It may be necessary to supplement word lists, since not all the sounds identified in the language will necessarily be found in all environments. Word lists for phonological analysis can be supplemented using the following elicitation techniques:

Syllable Prompts

Maddieson (2001:217) suggests that literate speakers can be very helpful in coming up with illustrative words. For example, the fieldworker could provide a written list of possible syllables, and ask if the native speaker can think of words that end or begin with these syllables. This task gives the native speaker input into the fieldwork process. It also allows the consultant to engage other members of the speech community in the fieldwork process, since s/he can use a written list to think about and discuss possible forms with other speakers away from the elicitation venue.

Substitution Tasks

Create words respecting the phonotactics of the language, and see if the consultant will accept and assign a meaning to the created words. This can be done in a number of ways. One example is (1), a sample dialogue from Abbi (2001:102–103):

(1) Fieldworker: Did you say the word for half is ad^{h}a?
Speaker: Yes.
Fieldworker: Do you also have a word such as at^{h}a?
Speaker: Not at^{h}a but we have ata.
Fieldworker: What does it mean?

When working in this fashion, there is a danger of creating a taboo word. Some native speakers won't be fazed by this at all and will readily provide a meaning for the "created" word. But others may be embarrassed by the taboo form and avoid assigning the word a meaning. Ladefoged (2003:6) recommends skipping an embarrassing word when it crops up in elicitation and proceeding without comment to the next task. A list including created words should be long enough to allow for the fact that some consultants may not recognize particular words on the list, or may not want to produce them.

Substitution with Rhyming

Another substitution task that is especially helpful when trying to contrast sounds in word-initial or word-final position is to create words that either rhyme or alliterate with words from the word list. If the concept of rhyming is alien to the native speaker, it may be difficult to implement this technique (Bowern 2008:38).

The procedure of organizing the data, developing a hypothesis of sound distribution, and checking this hypothesis with input from the native speaker is repeated until the fieldworker feels confident about suspected allophonic variation. How many times this has to be done depends on the language, the fieldworker's ear, and whether the fieldworker has done his or her typological homework. To end this rechecking phase, one might re-elicit, from different speakers, the words that led to the hypothesis about allophonic distribution.

10.8 Suprasegmentals

Some special comments need to be made about fieldwork on tone, stress, intonation, and other prosodic features. It is challenging to collect data on these aspects of the sound system because many factors influence their production. These include microprosodic effects, such as the lowering of fundamental frequency due to adjacent voiced consonants, and changes of duration, intensity, and pitch depending on the intention of the speaker. As with segmental analyses, the researcher's native language can affect how well suprasegmental features are perceived. In addition, some people seem to have a natural ability to hear pitch differences while others don't and this obviously will affect the interpretation of tone and intonation.

10.8.1 Fast Speech

In the advanced stages of phonetic and phonological fieldwork, the analysis of conversations and other texts can be used to discover fast speech phenomena. While all languages have some fast speech phenomena, some have more than others: for example, while Western Apache and Navajo (Athabascan) are closely related, Western Apache has more fast speech phenomena than Navajo. Of course, the best data for this comes from listening to naturally occurring speech, and identifying commonly recurring strings that are difficult to understand until they are repeated back slowly. The fieldworker could make a list of these unrecognizable and phonotactically strange strings and look for them in texts. Many native speakers will not recognize fast speech forms when discussed in isolation. For instance, the Manipuri fast speech form [hayšidətəw] from *hayribəsidə təw* 'thus did' was not recognized by my [Chelliah's] consultants until I showed it to them in a text.

10.8.2 Syllables and Phonotactic Constraints

Early on, it is advisable to make a list of the possible syllable shapes in the target language, and to use this information in formulating statements about sound distribution. Also, some phonological rules are based on syllable weight, so in these cases it will be necessary to know what counts as a light or heavy syllable. Vaux and Cooper (1999:88) suggest the following methodology to do this:

- Use words that have already been elicited to identify possible syllable initial and final clusters.
- When unsure of syllable breaks, ask for a speaker's judgment.
- Supplement decisions on syllable boundaries by having speakers say the words slowly to see where the pauses occur. But be aware that speakers can insert schwas to accentuate syllable breaks.
- Poetic rhyming copies the nucleus and coda to create rhymes, so rhymes can be used as clues to syllable boundaries. Even if native speakers don't know any poetry in particular, they may enjoy sharing language games or tunes they learned as children. These will have rhymes in them. Nicknames and other abbreviations also provide clues to syllable breaks.
- If the language has stress, there may be correlations between stress and syllable type.

10.8.3 Stress

Stress – the relative prominence of one syllable compared to its neighbors – can be realized as one, two, or all of the following features: vowel length, full versus reduced vowels, raised pitch, and raised intensity (sound energy). When first transcribing a language, the fieldworker might find that s/he can hear stress clearly, or not at all. This is because when the fieldworker's native language has stress, s/he will have an advantage in hearing stress; when the fieldworker's language does not have stress, s/he may experience "stress deafness". In addition, the primary cues for stress in the fieldworker's native language will influence what s/he concentrates on when identifying stress in the target language. If, for instance, fundamental frequency is a primary cue for stress in the fieldworker's native language, then s/he might interpret high fundamental frequency as an indication of stress in the target language. But in fact, in the target language, it might be vowel length which is actually the primary indication of stress (Nancy Caplow p.c.).

Keeping these concerns in mind, we have the following brilliant suggestion. If the fieldworker can hear stress, s/he should mark it! Speakers should not be asked to repeat words several times in order to check if the stresses have been written down correctly because such requests for repetition will lead to annoyed consultants and unnatural responses, with vowel length being one thing that gets unnaturally extended (Abbi 2001: 94).

The stress in collected word lists should then be compared by grouping together disyllabic words to investigate primary stress, and longer words to investigate secondary stress. Of course, a polysyllabic word might just have a series of primary stresses as in Siberian Yupik Eskimo, where every other syllable is stressed.

Healey (1975:166) suggests the following sorting protocol: Make separate lists of one, two, and three syllable words. Within those lists, group together words that have the same stress, vowel length, and pitch patterns. Classification can be further refined by grouping together words with the same consonant environments. There will be a residual category of words that don't seem to fit any of the lists; these can be set aside initially. Then, a native speaker should be recorded saying the words in a frame. The fieldworker must then use his or her typological knowledge of stress systems as a clue to the possible patterns that might occur in the target language. Useful overviews of stress systems are Hyman (2006), Dixon (2010a:280–283), and Goedemans et al. (2006).[7] Some patterns to look for are given below.

- Primary stress is fixed, based on the position of a syllable in the word, e.g. initial, ultimate, penultimate, or antepenultimate. For example, in Finnish word-initial syllables always receive primary stress.
- Primary stress is dependent on syllable weight, as in Latin, where stress falls on the penultimate syllable when that syllable is heavy. Otherwise, stress falls on the antepenult. Mongolian stresses the first heavy syllable if there is one, otherwise it stresses the first syllable. This example is from Vaux and Cooper (1999:88).
- Reduced vowels are not stressed.
- Clitics may or may not be unstressed. In languages such as Lakota, clitics are unstressed, so while they are phonologically part of the constituent they attach to, they are not seen by stress rules. In French, on the other hand, clitics are phonologically part of the constituent they attach to and are seen by stress rules.
- Just because a vowel is long, it is not necessarily stressed. Vowel length may be due to syllable closure or compensatory lengthening, for example. Similarly, vowel intensity may be due to vowel height, low vowels being intrinsically more intense than high vowels (Lehiste 1970).

[7] Hyman (2009) questions the usefulness of categorizing languages as either "stress" or "pitch accent" languages, arguing that (1) the label pitch accent only serves to tell us that a language has highly constrained pitch patterns. Languages categorized as pitch accent languages don't necessarily share other prosodic features, and in fact in some cases can look more like tone languages. (2) There are languages that are neither prototypical tone languages (one tone per syllable, one syllable per tone) nor prototypical pitch accent languages (one tone per word, one word per tone).

In terms of fieldwork, we suggest that the idea of a prototype, even if it is an idealized abstraction, is useful. Furthermore, just because there are non-prototypical languages does not mean that prototypical pitch accent languages do not exist. They do. If pitch accent is rigorously defined, Swedish, Norwegian, Limburg dialects of Dutch would qualify as prototypical examples. Manipuri – which mostly has one stem tone per word but also has suffixes with underlying tone – would not. Cantonese, which has no toneless syllables, is a prototypical tone language, while Mandarin Chinese is not since it can exhibit toneless syllables (Randy LaPolla p.c.).

Ladefoged (2003:92–94) explains the difficulties of instrumentally studying stress. First, because different sounds have different inherent intensity – voiced sounds are more intense than voiceless ones; vowels with lip rounding are more intense than spread vowels – intensity is not a good instrumental indication of stress. In fact, as stated above, stress is typically indicated by a combination of several factors in addition to intensity, i.e. pitch, vowel duration, and vowel quality. It is a language specific fact which combination of these resources is used.

Native speaker input is essential in figuring out stress in the early stages. The speaker could be asked to stamp, clap out, or use hand gestures to indicate the stress pattern of a word. Note that it is essential to combine work on stress with other tasks, since speaker fatigue can set in rather quickly. Once a list of words that seem to follow the same stress pattern have been compiled, these words can be re-elicited in frames such as this syntactic frame used by Nancy Caplow to study stress in Balti and Rebkong Amdo Tibetan: "In my language, we say _______" (Caplow 2009). Since Tibetan is a verb-final language, this frame placed the target word in utterance-medial position. Frames are necessary because – as with all phonetic data – the surrounding phonetic environments can change stress patterns, as can the pragmatics of the speech situation.

10.8.4 Tone

Attention to the following four factors will help with the challenges of investigating a tone language: selection of consultants for work on tone; systematic study of tonal data; practice hearing and reproducing tone; and development of a tone transcription system.

10.8.4.1 Selection of Consultants

A native speaker's tonal proficiency is comprised of the ability to:

- Produce clear tonal distinctions in natural and elicited speech
- Perceive tonal distinctions in meaningful contexts
- Repeat words with the proper tone out of context
- Provide consistent descriptions for tonal distinctions

Ideally, a speaker will have all four abilities, but since this is rarely true, the fieldworker will need a number of speakers to complete a tonal study. The first step is to find a speaker whose production of tone is clear.

Second, it is necessary to find a speaker who can identify the tone of a word. The skill of identifying the relative pitch of a syllable may be acquired with some training. The fieldworker could demonstrate how a tone can be whistled or hummed or played on an instrument such as a flute. Bouquiaux et al. (1992:17) advise that whistling may be taboo in some cultures and so, in those cases should not be used.

Note that not all speakers can be trained to hum or whistle a tone. Some will hum at high or low pitch in a seemingly random fashion. With other speakers the humming, whistling, or other indication of tone will correspond clearly with the tones being produced. Robert Rankin (p.c.) tells of one Chinese teacher who raised his eyebrows every time he produced a high tone. Consultants can be taught to tap on glasses which have different amounts of water in them to indicate if the syllable pronounced has a high, mid, or low tone. The glass prop is useful for level tones but not contour tones (Bouquiaux et al. 1992:19–21).

Note that second language learners may acquire tone but may not be able to identify it, or may downplay its importance; they might pronounce the tones correctly but report that they speak a toneless variety. Finally, it will also be necessary for consultants to discuss and name tones using their own or the fieldworker's metalanguage. Another issue is that speakers may use meaning as a guide to tone distinction, so they may identify words that are homophones as having two different tones.

10.8.4.2 Frames

In addition to the reasons for using frames in phonetic study in general, it is particularly important to study words in frames or carrier phrases – as well as in isolation – for tonal study. Adjacent consonants can lower or raise the pitch of vowels, and adjacent tones can assimilate, so the words under study should be recorded in frames that provide a consistent phonetic environment.

Additionally, in some languages, tone is hard to perceive. Some fieldworkers report that it can be easier to hear contrasts in a language with more phonemic tones than in a language with fewer phonemic tones; furthermore, contour tones are easier to perceive than level tones. Expectations also color perception. Randy LaPolla (p.c.) notes that Chinese linguists tend to describe languages as having tones even when they are toneless, while Indian scholars tend to describe languages as toneless even when they have tones. An unusual case of a linguist hearing tone where none exists is Whorf (1993): he found tone in Nahuatl, a Uto-Aztecan language, in which there is, in fact, no evidence for phonemic tone. One can surmise that this mis-analysis resulted from an overenthusiastic search for exotic linguistic features. In some cases, fluent speakers can produce the tonal distinctions but not hear them at all, and may resist the analyst's attempts at studying tone.

Sometimes tone is salient only in words of several syllables. In Manipuri, for example, rising and falling stem tones become increasingly easier to perceive as more suffixes are added to the word, because stem tone spreads to following syllables (Chelliah 1997). However, as Steve Morey (p.c.) pointed out to us, extending the length of a word with the same syllables is not always possible if, for example, the words of a minimal pair belong to different verb or noun classes.

One approach is to randomize the sequence of tones in the list of words to be recorded, so that, for example, all the rising tones are not together on the list. See Bouquiaux et al. (1992:14–15) on tone and frames for African languages.

The frame used will vary depending on whether or not the target language is verb final. It is best not to have the token under study at the end of a phrase, since phrase-final declination will affect the relative value of the tone on the phrase final syllable. Frames will also vary depending on the lexical category of the token. They will also vary depending on whether a minimal pair indicates lexical information (as in Manipuri *mí* 'man' versus *mi* 'spider') or grammatical information (as in the Nambikuára example, cited by Larry Hyman[8]: present tense *xyǎu-nàrǎ* 'he is staying' versus past tense *xyǎu-nǎrǎ* 'he stayed').

10.8.4.3 Transcription

Tones should be characterized and a transcription system established fairly early on in fieldwork to allow for consistent transcription. If a language family has been well described, it seems prudent to follow the transcription system already in place. If this is the first time a language is being transcribed, numeral or IPA transcriptions for tone are the easiest to use and keep consistent. Once the system is understood, accent marks or other diacritics can be used to simplify typing up data and to make it more readable.

Using numbers to indicate tone can be a useful training tool for native speakers who are helping with transcription. This is especially true if they can hear tones and reproduce them when asked, but have difficulty identifying them by category label. See discussion in Gudschinsky (1959). The numbers do not refer to absolute pitch values but are relative values for each speaker, so that 1 for a male speaker will be lower than 1 for a female speaker.

Narrow transcription should be used while studying tone until tone assimilation and spreading rules are discovered. See Yip (2002: 164–166) and Bao (1999: 61–75) for a detailed discussion of tonal phonology. In addition, differences in phonation often co-occur with tone distinctions so it is important to transcribe these features along with tones. In our view, a practical orthography should never replace tone transcription because orthographic conventions usually ignore interesting phonetic details and may not indicate tone at all. See Gudschinsky (1958) for more on practical orthographies and tone transcription. Of course, for purposes of publication or dissemination to the community, one can always produce a document using a standardized orthography based on the transcribed document.

10.8.4.4 Confirmation

Once a first pass has been made at transcribing words, the fieldworker should confirm or revise transcriptions as necessary. Have the right tones been associated

[8] http://linguistics.berkeley.edu/phonlab/annual_report/documents/2007/Hyman_Blackwell_Tone_PLAR.pdf.

with each word, and have the correct number of tones been identified? Here are some techniques to help with this checking phase:

1. Pitch tracking. It is fairly easy these days to instrumentally check on tone transcriptions. Thus we recommend that pitch tracking be used for confirmation and quantification; it should not be used before the fieldworker has formed initial hypotheses on the basis of auditory-based transcriptions and the input of the consultant. It is important to have some idea of tone patterns beforehand, so if the results of the instrumental study do not correspond with expectations, the fieldworker will know that there may have been a mistake in the set-up of the pitch tracking software, the way the instrumental study was set up, or the initial transcription. Pitch tracking can be conducted on site to facilitate the re-elicitation of words and improve on input data right away. Jerry Edmondson (p.c.) records words directly onto his computer and then runs the words through pitch tracking software in the field. This allows him to refine his analysis and prepare for further questioning of his speakers. See Ladefoged (2003:75–90) for a discussion of pitch tracking. A commonly used pitch tracking software at the time of writing is PRAAT. It can be downloaded free from www.praat.org and is available for the Mac and the PC. See Gordon (2003) for a discussion of other software possibilities for acoustic analysis including pitch tracking.
2. Native speaker input should be used for confirmation. In the initial stages of a study, speakers' comments on tone cannot be accepted without question. However, speaker input can be useful in controlled tasks such as producing minimal pairs and placing words in different tonal contour lists as described below:

   - Minimal pairs for tone study: Consultants should be given repeated exposure to minimal tone pairs. This will teach them to extend the list with more tone minimal pairs.
   - The native speaker can be taught to isolate the "odd tone out" using the MONOTONY TEST as described in Gleason (1961:302–304). This technique involves collecting several monosyllabic words, transcribed as best as possible. The words should be sorted into columns so that each column has a distinctive tone contour or pitch. Each column should be given a name like High, Low, or Mid. Then the words should be shuffled around so that they are in the "wrong" column. The consultant should then read each column out as a list. If the consultant is illiterate, the words can be re-elicited. Does any word stand out as not echoing the tonal melody of other words in the list? Determining this will help in reshuffling words to the appropriate list. Next the same test can be performed with disyllabic words. To avoid the effect of lowered pitch in list-final words, it is recommended that the order of items in a list be shuffled and the words re-elicited and rerecorded. Not all speakers can help with this test.
   - The fieldworker could show the native speaker how to associate a label or diacritic with a set of words based on a reference point. For example, the native speaker knows that the tone of words in Column A is called "rising" tone. A word under consideration has the same sound; therefore, it also goes in Column A and is labeled as rising.

- Another technique is to establish words for which the rhyme or tone information is known. The fieldworker can then use these words as “yardsticks” by which to judge other words by asking the consultants if the words in question are the same or different from the reference word (Randy LaPolla p.c.). Again, not all consultants will be able to judge one word out of context against another.

3. Fieldworker “mistakes” should be used as clues. The fieldworker uses his or her errors as a language learner to learn the correct tones of a word. Whether tonal errors are tolerated or not depends on the language and the native speaker community. In the case of Manipuri, there are many second language learners who do not acquire Manipuri tone. Native speakers tolerate the tonal errors as long as the learners can be understood. So, in Manipur at least, speakers will often tolerate mistakes or strange “accents” from the fieldworker.
4. Previous knowledge of the language family should be used for tone analysis: If comparative work has been done on a language and reconstructed forms for the language family or group are known, the fieldworker could use this information to guide tone transcription. Since sound change is regular, one would expect, for the most part, that target language tones will systematically correspond to reconstructed tones or derive from reconstructed phonation contrasts. This method has been used for the documentation of tone in many Athabascan languages (Krauss 2005). Hypotheses on tone based on comparative evidence would then be tested through native speaker perception tests or overt discussion with consultants who are able to hear and reproduce tone (Krauss 2005:131–133).
5. Perception tests can be used to confirm the number and types of phonemic tones in a language. An example of how to conduct and use the results of perception tests is given in Coupe (2003:95–99). First, an audio file with randomized words from tone minimal pairs, including some distracters, should be prepared using a variety of speakers if possible. Several tokens of the target words – words which exhibit the tonal contrasts under consideration – should be included (Coupe includes nine tokens of each target form). To avoid priming the respondent, the forms should not be recorded on the same day as the perception test. The perception test will entail playing back the audio file to a speaker and recording the speaker’s identification (meaning) of that form. The results of Coupe’s first perception test are presented in the confusion matrix in Table 10.1.

Table 10.1 Confusion matrix for four words of a tone minimal set (From Coupe 2003:96)

Stimulus → Response↓	/tə̄máŋ/ (MH)‘all’	/tə̄-māŋ/ (MM)‘NPF-body’	/tə̄-māŋ/ (LM)‘PROBH-believe’	/tə̄-màŋ/ (HL)‘NPF-dark’
/tə̄máŋ/ (MH)‘all’	9/9			
/tə̄-māŋ/ (MM)‘NPF-body’		9/9	4	
/tə̀-māŋ/ (LM)‘PROBH-believe’			5/9	
/tə́-màŋ/ (HL)‘NPF-dark’				9/9

NPF ‘nominalizing prefix’; PROBH ‘prohibitive’

From Table 10.1 we see that the speaker correctly recognized the Mid-High, Mid-Mid and High-Low tone sequences nine times out of nine (see gray shaded area). The Low-Mid is confused for the Mid-Mid four times out of nine (the number in the non shaded area). Further acoustic study explains the confusion: the pitch curves of the second syllable of the Low-Mid and Mid-Mid words are very similar. The main difference seems to be that the Low-Mid syllable has a slightly longer final syllable than the Mid-Mid. This experiment confirms for the fieldworker that some tones, while phonemic, may be hard to perceive and may be misreported by the speaker and misheard by the fieldworker.

6. Correlates of tone may be used as a guide. There are some common language specific correlations between phonation and tone that may serve as cues to tone. For example, breathy phonation is often associated with low tone, and creakiness is often associated with low or falling tone. However, some correlations may confound perception: a long vowel with a falling tone may be perceived as lower than the same short vowel with a falling tone. Vowel height also correlates with tone: high vowels are inherently higher in pitch than low vowels.

10.8.5 Intonation

Intonation is the manipulation of pitch across a phrase, clause, or sentence, usually in conjunction with other prosodic features such as segment duration, stress, and vowel quality. Unlike lexical tone, which is assigned in the lexicon and is an inherent part of a lexical item, intonational pitch differences are determined by communicative factors such as the intended illocutionary force of an utterance, and how a speaker packages new or presupposed information.

To begin the study of intonation, it should first be established whether or not the language has lexical tone (as in Cantonese, where each syllable has a lexically assigned tone) or pitch accent (as in Norwegian, where the placement of an accent determines the pitch contour of an utterance). The fieldworker must consider how to distinguish between intonation and tone and discover how intonation interacts with the existing lexical prosodic features of an utterance.

One way of determining if a word has lexical tone, or if the perceived pitch difference is due to intonation, is by moving the target constituent to different parts of a frame, e.g. *Ravi is coming to dinner. Is Ravi coming to dinner? I think the person who is coming to dinner is Ravi.* Unless the differences can be explained by tone assimilation rules, if the pitch changes on the target word, then phrase or clause level prosody is being observed. If the pitch remains more or less constant, lexical tone is being observed. Similarly, the target constituents could be studied in utterances reflecting a variety of contexts such as surprise, anger, confirmation, and questioning. If the pitch curve of a target word changes with each context, then the pitch does not reflect lexical tone but is due to the prosodic features indexing that context (Himmelmann and Ladd 2008:264).

Lists are useful for distinguishing pitch differences due to lexical tone and sentence level prosody. First, lists have predictable intonation – non-final tokens have rising intonation and final tokens have falling intonation. Second, lists are not usually affected by pragmatic concerns like focus or topic. See Schiffrin (1994:291–317) and Himmelmann and Ladd (2008:264) for more on lists. So, for example, one could switch around the order of words in citation forms in a list e.g. *Jane, Jill, Janice; Janice, Jane, Jill; Jill, Janice, Jane.* If the pitch remains the same regardless of the position of the word in the list, lexical tone is indicated (Himmelmann 2006:176).

We know that there is not a one-to-one correspondence between intonation patterns and sentence types. For example, in English it is impossible to say what pattern will occur with a declarative sentence without knowing how the sentence is used. If used to relay information, a declarative will probably have a phrase final fall in pitch – but this is not true when a declarative is used to issue a warning, to tease, or to scold. Another example is the wh-question in English, which has different contours depending on whether the question is polite or neutral with respect to politeness (Himmelmann and Ladd 2008:264). Furthermore, intonation can be modulated depending on a speaker's emotional state and the interlocutor's reception of the utterance, e.g. I might have to scream out a warning if the hearer is day-dreaming.

Given that these factors affect intonation, how best to devise elicitation schemes to document possible intonation contours? Himmelmann (2006:167–168) suggests having consultants, "produce the same utterance in the same context with the same intention" by narrating the same story. He gives the example of comparing different retellings of the folktale "Little Red Riding Hood", where particular phrases, such as "What big ears you have!" are expected to be uttered. Retellings of the Pear Story, or jokes or traditional stories can be used as well. Another task Himmelmann (2006:170) suggests is the "map task", where a speaker asks questions about where something is located and the interlocutor answers. Similarly, if the object of study is declarative intonation, the speaker can ask identification questions about objects such as "What is that?" to elicit a declarative response. All of this reinforces the benefits of collecting the same texts from a variety of speakers.

For other controlled elicitation tasks, a fieldworker can construct conversational adjacency pairs such as question–answer, warning–acceptance, polite request–acquiescence, command–agreement, or compliment–reaction. See Himmelmann (2006:170–171) for a set of examples. If the consultants are literate, the pairs could be written down and the consultants could read these back. Some speakers do a fantastic job of creating or reading dialogues in a natural way, but this is a special talent and cannot be expected of everyone. It helps greatly to review an interactional script with a speaker who is a good "actor" and to remedy any awkwardness in the phrasing. These corrected scripts can then be used with other speakers to elicit the adjacency pairs. Some speakers won't understand this task; they will not be able to use the scripts to act out an imagined interaction, but instead will talk around the examples or simply read them aloud. This does not mean that the task is bad, just that it is not appropriate for these particular speakers.

With endangered languages, even if the language in question does have a writing system, it is likely that most consultants will not know how to read and write fluently, so reading will be unnatural (Himmelmann and Ladd 2008:66). The ideal situation is to have two speakers who can engage in natural conversation. The vocabulary used in these conversations can be controlled by giving the speakers a specific topic, picture prompt, or other goal-oriented interaction such as giving or getting directions. Himmelmann and Ladd (2008:266) provide references to tasks that can be used for such elicitation.

Perhaps the most accurate understanding of intonation comes from direct observation of natural interactions. The fieldworker should try to observe, mimic, and make notes of intonational contours in commands, statements, questions of various types, and focus constructions such as clefts, as they occur in conversations, personal or traditional narratives, exhortations, arguments, expressions of anger, scolding, teasing and ridiculing, baby-talk, and complaints. See Healey (1975:37–40) for a longer list of speech acts, and Lee et al. (2007) for more on intonation and focus. The fieldworker should make notes on the intonational contour most often associated with different sentence types and use this as a basis for more directed elicitation.

Apart from increasing our general knowledge of how intonation is used in the world's languages, analyzing intonation patterns has practical applications for linguistic fieldwork. When transcribing natural speech, very often the only clue the fieldworker, and even the speaker, have about constituency come from intonation. It would be helpful to have some knowledge of how speech acts correspond with intonation contours to make good decisions about constituent analysis and text translation.

When recording utterances for analysis of intonation contours, the same principles used for recording tone should be followed. High quality recordings should be made to allow for acoustic analysis. Utterances should be recorded so that they can be compared. This means that the recorded tokens should contain similar lexical items, minimizing the chance of lexical tone, vowel quality, word length, or other microprosodic factors contributing to a difference in intonational contour. When constructing elicitation schedules, Himmelmann (2006:171) recommends avoiding lexical items with voiceless stops and fricatives because continuous voicing is needed for pitch tracking.

Several speakers of both genders should be recorded. Women often have greater intonational variation than men, and younger speakers may be more animated than older speakers. (See Himmelmann and Ladd 2008:265). Himmelmann (2006:169) recommends recording at least two male and two female speakers, and collecting at least four to ten tokens of the same utterance, preferably produced at different times. For comparability, the minimum required is three tokens of the same utterance.

Finally, when working with a tone language, the fieldworker must decide on a transcription system for intonation. If a purely iconic system – such as a line over an utterance is used – the fieldworker should be careful to keep the transcription of tone and intonation distinct. For example, notations should not conflate falling tone with the indication of declination. Transcription systems and software for intonation are available such as ToBI (Tones and Break Indices). The ToBI software makes it possible to build online databases of prosodically transcribed natural speech with standardized annotation conventions for different language families (Beckman et al. 2005). See also Bolinger (1987, 1989) for a discussion of the International Transcription System of Intonation.

10.9 Conclusion

Today's phonetic fieldwork requires the use of sound analysis software. The fieldworker must be proficient in basic skills such as reading wideband spectrograms to segment an utterance into discrete phones. Phonetic fieldwork may require the use of sophisticated hardware, as well. In addition to the use of technology, the fieldworker needs to respect the old staples of phonetic fieldwork: ear work, elicitation, transcription, and record-keeping. Phonetic fieldwork goes hand-in-hand with phonological analysis; at an early stage, the fieldworker should have hypotheses on the following: phonemes and allophonic variation; syllable structure; phonotactics; co-articulation; phonological processes; stress rules and tone rules.

References

Abbi, Anvita. 2001. *A Manual of Linguistic Field Work and Indian Language Structures.* (Lincom Handbooks in Linguistics 17.) Munich: Lincom Europa.

Arka, I. Wayan. n.d. Palatography in a fieldwork setting: investigating and analyzing alveolar continuant [r] and [ɹ] in Rongga. Online: http://rspas.anu.edu.au/linguistics/projects/iwa/Web-Pages/PalatographyRongga%20–%20Wayan%20Arka.pdf

Austin, Peter K. 2006. Data and language documentation. In *Essentials of Language Documentation*, ed. by Jost Gippert, Nikolaus Himmelmann, and Ulrike Mosel, 87–112. Berlin and New York: Mouton de Gruyter.

Bao, Zhiming. 1999. *The Structure of Tone.* Oxford: Oxford University Press.

Beckman, Mary E., Julia Hirschberg, and Stefanie Shattuck-Hufnagel. 2005. The original ToBI system and the evolution of the ToBI framework. In *Prosodic Models and Transcription: Towards Prosodic Typology*, ed. by Sun-Ah Jun, 9–54. Oxford: Oxford University Press.

Bolinger, Dwight. 1987. *Intonation and Its Parts: Melody in Spoken English.* Stanford, CA: Stanford University Press.

Bolinger, Dwight. 1989. *Intonation and Its Uses: Melody in Grammar and Discourse*. Stanford, CA: Stanford University Press.

Bouquiaux, Luc, Jacqueline M. C. Thomas, and James Robert. 1992. *Studying and Describing Unwritten Languages.* Dallas, TX: SIL International.

Bowern, Claire. 2008. *Linguistic Fieldwork. A Practical Guide*. New York: Palgrave MacMillan.

Caplow, Nancy J. 2009. The role of stress in Tibetan tonogenesis: a study in historical comparative acoustics. Ph.D. Dissertation, University of California, Santa Barbara.

Chelliah, Shobhana L. 1997. *A Grammar of Meithei*. (Mouton Grammar Library 17). Berlin and New York: Mouton de Gruyter.

Coupe, Alec R. 2003. *A Phonetic and Phonological Description of Ao. A Tibeto-Burman of Nagaland, north-east India* (Pacific Linguistics 543.) Canberra: Pacific Linguistics. Research School of Pacific and Asian Studies, The Australian National University.

Dixon, R. M. W. 2010a. *Basic Linguistic Theory. Volume 1 Methodology*. Oxford: Oxford University Press.

Dixon, R. M. W. 2010b. *Basic Linguistic Theory. Volume 2 Grammatical Topics*. Oxford: Oxford University Press.

Dixon, R. M. W., and Alexandra Y. Aikhenvald, eds. 2002. *Word. A Cross-Linguistic Typology*. Cambridge: Cambridge University Press.

E-MELD. 2005. *E-MELD School of Best Practice: Ultrasound Analysis: Navajo*.Online: http://emeld.org/school/case/navajo/index.html

Ferguson, Charles A. 1960. The phonemes of Bengali. *Language* 36(1):22–59.

Gick, Brian. 2002. The use of ultrasound for linguistic phonetic fieldwork. *Journal of the International Phonetic Association*, 32:2:113–121.

Gleason, Henry A. Jr. 1961. *An Introduction to Descriptive Linguistics*. New York: Holt, Rinehart & Winston.

Goedemans, Ron, Harry van der Hulst, and Ellis Visch. 2006. StressTyp. The Structure of a Database for Stress Patterns of the World's Languages. University of Leiden. Internet Version last updated 2/8/2006. Online: http://www.unileiden.net/stresstyp/manual.htm.

Gordon, Matthew. 2003. Collecting phonetic data on endangered languages. *15th International Congress of Phonetic Sciences*, 207–210.

Gudschinsky, Sarah C. 1958. Native reactions to tones and words in Mazatec. *Word* 14(2–3):338–345.

Gudschinsky, Sarah C. 1959. Toneme Representation in Mazatec Orthography. *Word* 15(3): 446–452.

Harris, Zellig S. 1951. *Methods in Structural Linguistics*. Chicago, IL: University of Chicago Press.

Haugen, Einar I. 1958. The phonemes of modern Icelandic. *Language* 34:55–88.

Healey, Alan. 1964. *Handling Unsophisticated Linguistic Informants*. (Linguistic Circle of Canberra Publications, Series A, Occasional Papers No. 2.) Canberra: Linguistics, Department of Anthropology and Sociology, The Australian National University.

Healey, Alan, ed. 1975. *Language Learner's Field Guide*. Ukarumpa, Papua New Guinea: Summer Institute of Linguistics.

Hildebrandt, Kristine. 2005. A phonetic analysis of Manange segmental and suprasegmental properties. *Linguistics of the Tibeto-Burman Area* 21.8: 1–36.

Hildebrandt, Kristine A. 2007. Phonology and Fieldwork in Nepal: Problems and Potentials. In *Proceedings of the Conference on Language Documentation and Linguistic Theory* ed. by Peter Austin, Oliver Bond and David Nathan, 33–44. London: SOAS.

Himmelmann, Nikolaus P. 2006. Prosody in language documentation. In *Essentials of Language Documentation*. (Trends in Linguistics. Studies and Monographs 178.), ed. by Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel, 163–181. Berlin and New York: Mouton de Gruyter.

Himmelmann, Nikolaus P., and D. Robert Ladd. 2008. Prosodic Description: An Introduction for Fieldworkers. *Language Documentation and Conservation* 2(2):244–274. Online: http://nflrc.hawaii.edu/ldc.

Hoijer, Harry. 1958. Native Reaction as a Criterion in Linguistic Analysis. In *Proceedings of the Eighth International Congress of Linguists*, ed. by Eva Sivertsen, 573–583. Oslo: Oslo University Press.

Hyman, Larry. 2006. Word-prosodic typology. *Phonology* 23: 225–257.

Hyman, Larry M. 2009. How (not) to do phonological typology: the case of pitch-accent. *Language Sciences* 31(2–3):213–238.

Krauss, Michael E. 2005. Athabascan Tone. In *Athabascan Prosody*, ed. by Sharon Hargus and Keren Rice, 55–136. (Current Issues in Linguistic Theory 269.) Amsterdam and Philadelphia: John Benjamins.

Krishnamurti, Bhadriraju. 2007. Fieldwork on Konda, a Dravidian language. *Sprachtypologie und Universalienforschung* (Focus on Linguistic Fieldwork, ed. by Alexandra Aikhenvald) 60(1):56–66.

Ladefoged, Peter. 2003. *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*. Malden, MA: Blackwell.

Ladefoged, Peter. 2005a. A Course in Phonetics [5th edition]. Boston, MA: Cengage Learning/Thomson/Wadsworth Publishers.

Ladefoged, Peter. 2005b. *Vowels and Consonants* [2nd edition]. Boston, MA, Oxford, Victoria: Blackwell.

Laver, John. 1994. *Principles of Phonetics.* (Cambridge textbooks in Linguistics). Cambridge: Cambridge University Press.

Lee, Chungmin, Matthew Gordon, and Daniel Büring (eds.). 2007. *Topic and Focus. Cross -Linguistic Perspectives on Meaning and Intonation Series: Studies in Linguistics and Philosophy*, Vol. 82. Dordrecht, The Netherlands: Springer.

Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, MA: MIT Press.

Maddieson, Ian. 2001. Phonetic fieldwork. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 211–229. Cambridge: Cambridge University Press.

Matthews, G. Hubert. 1955. A phonemic analysis of a Dakota dialect. *International Journal of American Linguistics* 21:56–59.

Redden, James E. 1982. On expanding the meaning of applied linguistics: A suggestion for training linguists and language teachers in field linguistics. *Notes on Linguistics* 22:19–23.

Schiffrin, Deborah. 1994. *Approaches to Discourse.* (Blackwell Textbooks in Linguistics). Oxford and Cambridge: Blackwell.

Sutton, Peter, and Michael Walsh. 1979. *Revised Linguistic Fieldwork Manual for Australia*. Canberra: Australian Institute of Aboriginal Studies.

The UCLA Phonetics Lab Archive. 2007. Los Angeles, CA: UCLA Department of Linguistics. Online: http://archive.phonetics.ucla.edu

Vaux, Bert, and Justin Cooper. 1999. *Introduction to Linguistic Field Methods*. (Lincom Course books in Linguistics 1.) Munich: Lincom Europa.

Whorf, Benjamin L., ed. by Lyle Campbell, and Frances Karttunen. 1993. Pitch Tone and the "Saltillo" in Modern and Ancient Nahuatl. *International Journal of American Linguistics* 59(2):165–223.

Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.

Chapter 11
What to Expect in Morphosyntactic Typology and Terminology

It seems, however, better not to complicate our present discussion by taking into consideration the possibilities of exact expression that may be required in idiomatic forms of speech, but rather to consider only those parts of the sentence which according to the morphology of the language, must be expressed. (Boas 1911:43 [emphasis in the original])

(…) if one looks for a particular category in a language one will usually find it: early European grammarians found pluperfect subjunctives in languages the world over. But then you have to ask yourself: how would I have interpreted the grammar of this language if English had never existed?" (Halliday 1994:xxxiii)

11.1 Introduction

There is no single right way of doing fieldwork, but there are some wrong ways. For example, it would be wrong for a fieldworker to begin a data gathering session with no knowledge of the possible grammatical structures that s/he is about to encounter.

In this chapter, we provide guidelines – albeit of a sketchy nature – to assist the beginning fieldworker in isolating the morphosyntactic, morphological, and syntactic characteristics of the target language, and we provide terminology that will help in labeling and identifying these characteristics. In order to succeed, a fieldworker must also know effective data gathering methods. These are discussed in Chapter 12.

This chapter is based on the premise that data gathering and data analysis cannot be carried out in sequence. To be sure, one starts with a bit of data gathering, but as soon as that bit of data has been collected, analysis is consciously or unconsciously carried out; this analysis stimulates more data gathering which results in further analysis, which stimulates further data gathering, and so on. Each stage of analysis prompts an evaluation of the data collected. Were sufficient and appropriate examples recorded? Were transcriptions and translations correct? Each set of data informs and advances the analysis.

Thus, as Nida (1949:192) put it, "Analysis must go hand in hand with fieldwork." In the crime shows we see on television these days, the police detective

S.L. Chelliah and W.J. de Reuse, *Handbook of Descriptive Linguistic Fieldwork*,
DOI 10.1007/978-90-481-9026-3_11, © Springer Science+Business Media B.V. 2011

gathers the evidence in a neat little plastic bag, which s/he hands over to a technician who analyzes the evidence, and the analysis is then handed over to a computational criminologist who generates a full description of the crime. It would certainly be interesting to be able to ask the fieldworker to gather the language data in a neat digital package that could be handed over to a linguist at headquarters who would do the analysis, or, as some wish, to hand over the analysis to a computational linguist to generate a full description of the grammar of the language. But linguistic fieldwork is messier than police detective work: it is impossible to neatly separate data collection (the documentation) from data analysis (the description), as was argued already in Section 2.2.1.3.

So the fieldworker must recognize that data collection and analysis go hand-in-hand; for every bit of data collection and documentation, there is some equal amount of analysis and description that occurs. And the fieldworker must be ready and able to formulate, test, and revise hypotheses as data collection proceeds.

In Section 11.2, we discuss grammatical typology and terminology in general, focusing on what is relevant to morphology and syntax. In Section 11.3, we focus on morphological typology and terminology. In Section 11.4, we focus on syntactic typology and terminology, and we offer concluding remarks in Section 11.5.

11.2 Grammatical Typology and Terminology

11.2.1 Preparatory Reading for Grammatical Fieldwork

To prepare for grammatical fieldwork, we cannot emphasize enough the importance of either taking a Linguistic Typology course or reading some general typological surveys and grammars. The most useful books on linguistic typology are those that discuss a wide range of constructions from typologically diverse languages. This reading can then be followed by more specific investigation into the structures of languages in the family being considered by the fieldworker. We suggest this general-to-specific progression in preparatory reading because it will prevent the proverbial language family “blinders” from coming on at an early stage. If data is collected with knowledge of only a limited set of languages or language families, the descriptive traditions of that family may hamper innovative analyses.

11.2.1.1 Typological Surveys and Manuals

For engaging introductions to morphosyntactic typology appropriate for the undergraduate student interested in fieldwork, we recommend Bird and Shopen (1979) on Maninka of West Africa, Haiman (1979) on Hua from Papua New Guinea, Haviland (1979) on Guugu Yimidhirr of Australia, Keenan and Ochs (1979) on Malagasy of Madagascar, Zúñiga (2007) on Mapudungun of Chile (with many typological comparisons with other languages), Abbi (2001) on the languages of

India, Lynch (1998:100–202) on Pacific languages (including Papuan and Australian languages), Silver and Miller (1997:17–45), on Native American Languages, Welmers (1973) on the languages of Africa, and Lyovin (1997) on the languages of the world.

More advanced surveys can be used as guides to morphosyntactic fieldwork. Peck (1981), Mallinson and Blake (1981), and Hagège (1982) are earlier surveys. Payne (1997) is the best one-volume guide, and provides a large set of morphological and syntactic terms. In this way it is a useful consciousness-raiser. The Whorfian notion that if you have a name for something, you will be able to identify it more easily is certainly applicable when trying to describe a language. This volume is also useful in that, due to its functionalist orientation, it recognizes the importance of prototypes and fuzzy categories.

More typologically oriented advanced introductions are Song (2001) and Whaley (1997).

Song (2001) is focused on the Greenbergian tradition of typology, and relies heavily on secondary sources. Only four typological issues are treated: basic word order, case marking, relative clauses, and causatives. The book ends with two interesting chapters, one on the applications of linguistic typology, and one on three European approaches to linguistic typology (Leningrad, Cologne, and the Prague School).

Whaley (1997) includes a history of typology and discusses issues of method and explanation, but a central focus is the illustration of basic linguistic categories and structures, including information on constituent order typology (traditionally called "word order typology"), morphological types, relational and semantic properties of nouns, verbal categories such as tense and aspect, mood, negation, the encoding of speech acts, and complex clauses.

A comprehensive survey of morphosyntactic typology is the two-volume work in French by Creissels (2006a, b) which, even though it mentions only 'syntaxe' in the title, covers quite a bit of morphology as well. One of its appealing features is an unusually detailed language index, which enables the reader to quickly find sections with typological discussion of a particular language.

The most advanced textbook on morphosyntactic typology is Croft (2003). A bibliography on language typology is in Payne (1991).

The three volume collection entitled *Language Typology and Syntactic Description*, edited by Shopen (1985), has a substantially revised second edition (2007) and is in a class by itself. About half its content consists of new articles. This work provides a clearer link between typological issues and language description than either Whaley (1997) or Croft (2003). Notwithstanding the title, this collection covers morphology quite well. These volumes remind us that not every instance or subtype of a discussed construction can be found in the language a fieldworker is studying. For example, reading about complementation (Noonan 1985, 2007) in the Shopen volumes can serve as a reminder of relevant terminology, of the existence of different kinds of complementation (see also Section 11.4.5), and of common relationships between complementizers, main clauses and subordinate clauses. Thus reading Noonan does not tell one what to look for per se, but rather, as a good guide should, it helps the fieldworker isolate and identify known constructions, and compare these to new constructions.

Typological surveys of particular grammatical constructions can also be useful as an advanced guide. Examples are Aikhenvald (2000) on classifiers, Haspelmath (1997) on indefinite pronouns, and Peterson (2007) on applicative constructions.

The most recent typological survey – and one of the most comprehensive – is the two-volume work by Dixon (2010a, b), with a third volume to appear. This is an encyclopedic overview, boldly synthesizing the ideas about typology and the methodology of grammar writing that Dixon has held and published on for the last 20 years or so. The title of these volumes, "Basic Linguistic Theory", a term introduced in the short but influential Dixon (1997), is potentially somewhat misleading. Basic linguistic theory (henceforth BLT) seems to mean two separate things. On the one hand, BLT can refer to the approach to grammar based on semantic principles that Dixon has used in his grammar of English (Dixon 1991); see Chapters 2 and 3 of Dixon (2010a) in particular. This is a very commonsensical theory, but not everyone would consider it a complete theory of grammar. On the other hand, BLT can also refer to the common canon of terminology and basic linguistic knowledge which every modern linguist, regardless of theoretical orientation, will consider his or her own. Thus there is a lot of very basic information in this volume, and much of the book can be read by undergraduates. However, what beginning fieldworkers and grammar writers must be aware of is that the two conceptions of BLT are not always sufficiently distinguished. Nevertheless, though it is somewhat personal and occasionally repetitious in its emphases, there is much in this monumental and very well-written work that we completely agree with.[1]

Manuals that include exercises in morphosyntactic analysis tend to be less comprehensive, but are good for practice before going into the field.

Beginning manuals with exercises are Elson and Pickett (1988), its matching exercise book (Merrifield et al. 2003), and Lockwood (2002). Lockwood has an eclectic theoretical background (stratificational, tagmemic, systemic) which not everyone will find helpful, but it has many interesting exercises.

A more advanced manual for methods of morphosyntactic analysis is Bickford (1998). Bickford also has useful annotated bibliographical notes, and lists linguistic terms at the end of each chapter.

11.2.1.2 Descriptive Grammars

Some of the typological manuals such as Payne (1997:372–375), Chapters 24 and 25 of Bickford (1998), and Dixon (2010a, b) discuss the importance of grammatical description and grammar writing. When training for linguistic fieldwork, reading surveys and manuals is not enough, and nothing replaces the reading of linguistic grammars and grammatical sketches. Grammar writing is a complicated task which requires the researcher to "work on hundreds of different problems, in parallel mode, and keep track of [his/her] analytical decisions about each of them" (Evans 2008:3). Considering

[1] However, most linguists will not agree, for example, with Dixon's choice of words in making a distinction between "formal linguists" and "scientific linguists" (2010a:183). This is not the place to discuss this choice of words, and we trust that field linguists will overlook the controversies that such words might engender.

beforehand how these pieces fit together can help guide the investigative process, which also requires constant synthesizing of information from each grammatical area.

A useful and concise survey of types of grammars, and their uses and formats, is in Mosel (2006). Evans (2008:5–6) provides this list for preparatory grammar reading (among other materials): (1) grammars of the fieldworker's native language; (2) grammars of languages that the fieldworker knows well; (3) grammars of typologically diverse languages that the fieldworker is unfamiliar with; and (4) grammars of languages related to the one the fieldworker is studying. We would like to add that it is also useful to read grammars by native speakers, which are written with a combined reliance on introspection and fieldwork. Examples of these are:

Bamgbose, Ayo. 1966. *A Grammar of Yoruba*. (West African Language Monograph Series 5). Cambridge: Cambridge University Press.
McCawley, James D. 1998. *The Syntactic Phenomena of English*, 2nd ed. Chicago: Chicago University Press.
Saeed, John I. 1993. *Somali Reference Grammar*. (Second Revised Edition.) Kensington, MD: Dunwoody Press.

It is also interesting to pay special attention to grammars written by a fieldworker-native speaker linguist team. Examples of these are:

Bauer, Winifred, with William Parker and Te Kareongawai Ewans. 1997. *The Reed Reference Grammar of Maori*. Auckland: Reed.[2]
Maganga, Clement, and Thilo C. Schadeberg. 1992. *Kinyamwezi: Grammar, Texts, vocabulary*. Cologne: Rüdiger Köppe.

An areally classified list of other reliable fieldwork-based grammar models is provided below. This list consists of our own recommendations as well as recommendations found in Payne (1997:372–375), Good (2006), Rice (2007), and Evans (2008:5–8). Another useful source of recommended grammars, with commentary, and largely overlapping the list below, is in Dixon (2010a:81–85).

Here are some observations on this list. A few (numbers 6, 25, 28, 31, 63) are grammars from the North Holland or Croom Helm grammar series, now published by Routledge. These rigorously follow the Comrie and Smith (1977) questionnaire.

One grammar, 4, is organized alphabetically by topic, like an encyclopedia.

Grammars 3 and 22 straddle the line between reference grammars and pedagogical grammars.

Grammars 4, 34, 41, and 60 are exemplary grammars of languages with established written traditions, but which are nevertheless based heavily on fieldwork.

Grammars 9 and 15 are detailed in their coverage of phonetics, phonology, and morphology, but do not contain any syntax.

Some of these grammars are not easy to follow – for example 10 and 40 – mainly because no interlinear glossing is given. We believe, however, that the lack of interlinear glossing, while making the job of the reader more difficult, has no bearing on the quality of the grammar.

[2] This grammar is partially based on Bauer et al. (1993), a grammar of the Routledge Grammar Series.

Grammar 14 is one of the best grammars in the Boas–Sapir–Bloomfield tradition, a reminder that some of the older grammars are still unsurpassed and easy to read.

A useful bibliographical list of grammars published within the last decade or so is available on-line at the ALT Grammar Watch site (http://www.eva.mpg.de/lingua/tools-at-lingboard/pdf/grammarwatch.pdf).

Africa

1. Childs, G. Tucker. 1995. *A Grammar of Kisi: A Southern Atlantic Language*. (Mouton Grammar Library 16.) Berlin and New York: Mouton de Gruyter.
2. Dimmendaal, Gerrit J. 1983. *The Turkana Language*. Dordrecht: Foris.
3. Hutchison, John P. 1981. *A Reference Grammar of the Kanuri Language*. Madison, WI: African Studies Program, University of Wisconsin–Madison.
4. Newman, Paul. 2000. *The Hausa Language. An Encyclopedic Reference Gramm*ar. New Haven, CT: Yale University Press.
5. Noonan. Michael. 1992. *A Grammar of Lango*. (Mouton Grammar Library 7.) Berlin and New York: Mouton de Gruyter.

America, Creoles

6. Huttar, George L., and Mary L. Huttar. 1994. *Ndyuka*. London: Routledge.
7. Kouwenberg, Silvia. 1994. *A Grammar of Berbice Dutch Creole*. (Mouton Grammar Library 12.) Berlin and New York: Mouton de Gruyter.

America, Mexico and Central

8. England, Nora C. 1983. *A Grammar of Mam, a Mayan Language*. Austin: University of Texas Press.
9. MacKay, Carolyn J. 1999. *A Grammar of Misantla Totonac*. Salt Lake City: University of Utah Press.

America, North

10. Bergsland, Knut. 1997. *Aleut Grammar. A Descriptive Reference Grammar of the Aleutian, Pribilof, and Commander Islands Aleut Language*. (Alaska Native Language Center Research Paper No. 10.) Fairbanks, AK: Alaska Native Language Center, University of Alaska.
11. Davis, Philip, and Ross Saunders. 1997. *A Grammar of Bella Coola*. (University of Montana Occasional papers in Linguistics 13.) Missoula: University of Montana.
12. van Eijk, Jan. 1997. *The Lillooet Language: Phonology, Morphology, Syntax.* Vancouver: University of British Columbia Press.
13. Graczyk, Randolph. 2007. *A Grammar of Crow*. Lincoln: University of Nebraska Press.
14. Haas, Mary R. 1941. *Tunica*. New York: J. J. Augustin.
15. Hargus, Sharon. 2007. *Witsuwit'en Grammar. Phonetics, Phonology, Morphology*. Seattle, WA: University of Washington Press.
16. Hill, Jane H. 2005. *A Grammar of Cupeño*. (University of California Publications in Linguistics 136.) Berkeley and Los Angeles: University of California Press. Online: http://repositories.cdlib.org/ucpress/ucpl/vol_136

17. Kimball, Geoffrey D. 1991. *Koasati Grammar*. Lincoln and London: University of Nebraska Press.
18. Quintero, Carolyn. 2004. *Osage Grammar*. Lincoln and London: University of Nebraska Press.
19. Rice, Keren. 1989. *A Grammar of Slave*. (Mouton Grammar Library 5.) Berlin and New York: Mouton de Gruyter.
20. Suttles, Wayne. 2004. *Musqueam Reference Grammar*. Vancouver: University of British Columbia Press.
21. Thompson, Laurence C. 1992. *The Thompson Language*. University of Montana Occasional Papers in Linguistics 8. Missoula: University of Montana.
22. Valentine, J. Randolph. 2001. *Nishnaabemwin Reference Grammar*. Toronto: University of Toronto Press.
23. Watkins, Laurel J., with Parker McKenzie. 1984. *A Grammar of Kiowa*. Lincoln and London: University of Nebraska Press.

America, South: Amazonia

24. Aikhenvald, Alexandra Y. 2003. *A Grammar of Tariana, from Northwest Amazonia*. (Cambridge Grammatical Descriptions.) Cambridge: Cambridge University Press.
25. Derbyshire, Desmond C. 1979. *Hixkariana*. London: Croom Helm.[3]
26. Dixon, R. M. W. 2004. *The Jarawara Language of Southern Amazonia*. Oxford: Oxford University Press.

America, South: Andes

27. Adelaar, Willem F. H. 1977. *Tarma Quechua: Grammar, Texts, Dictionary*. Lisse: The Peter de Ridder Press.
28. Cole, Peter. 1982. *Imbabura Quechua*. London and Dover, N.H.: Croom Helm.
29. Weber, David J. 1989. *A Grammar of Huallaga (Huanuco) Quechua*. (University of California Publications in Linguistics 112.) Berkeley and Los Angeles: University of California Press.

Asia: Caucasia[4]

30. Haspelmath, Martin. 1993. *A Grammar of Lezgian*. (Mouton Grammar Library 9.) Berlin and New York: Mouton de Gruyter.
31. Hewitt, B. G., and Z. K. Khiba. 1989. *Abkhaz*. (Croom Helm Descriptive Grammar Series.) London and New York: Routledge. [1979, London: North-Holland]

Asia: East and Central Asia

32. Slater, Keith W. 2003. *A Grammar of Mangghuer. A Mongolic language of China's Qinghai-Gansu Sprachbund*. London and New York: RoutledgeCurzon.
33. Tamura, Suzuko. 2000. *The Ainu language*. Tokyo: Sanseido.

[3] The importance of Hixkariana for typology is emphasized further in Derbyshire (1985).

[4] Grammars for East Caucasian are discussed in Schulze (2005).

Asia: Middle East

34. Cowell, Mark W. 1964. *A Reference Grammar of Syrian Arabic (Based on the dialect of Damascus)*. Washington D.C.: Georgetown University Press.

Asia: South Asia

35. Coupe, Alec R. 2007. *A Grammar of Mongsen Ao*. (Mouton Grammar Library 39.) Berlin and New York: Mouton de Gruyter.
36. Genetti, Carol. 2007. *A Grammar of Dolakha Newar*. (Mouton Grammar Library 40.) Berlin and New York: Mouton de Gruyter.
37. Osada, Toshiki. 1992. *A Reference Grammar of Mundari*. Tokyo: Institute for the Languages and Cultures of Asia and Africa.
38. Watters, David E. 2002. *A Grammar of Kham*. (Cambridge Grammatical Descriptions.) Cambridge: Cambridge University Press.

Asia: Southeast Asia

39. Kruspe, Nicole. 2003. *A Grammar of Semelai*. (Cambridge Grammatical Descriptions.) Cambridge: Cambridge University Press.
40. Matisoff, James A. 1973. *The Grammar of Lahu*. (University of California Publications in Linguistics 75.) Berkeley and Los Angeles: University of California Press.
41. Thompson, Laurence C. 1987. *A Vietnamese Reference Grammar*. Honolulu: University of Hawai'i Press. [Originally 1965]

Australia

42. Austin, Peter K. 1981. *A Grammar of Diyari, South Australia*. (Cambridge Studies in Linguistics 32.) Cambridge: Cambridge University Press.
43. Dixon, R. M. W. 1972. *The Dyirbal Language of North Queensland*. Cambridge: Cambridge University Press.
44. Dixon, R. M. W. 1977. *A Grammar of Yidiŋ*. Cambridge: Cambridge University Press.
45. Donaldson, Tamsin. 1980. *Ngiyambaa, the Language of the Wangaaybuwan*. Cambridge: Cambridge University Press.
46. Evans, Nicholas. 1995. *A Grammar of Kayardild, with Historical-Comparative Notes on Tangkic*. (Mouton Grammar Library 15.) Berlin and New York: Mouton de Gruyter.
47. Heath, Jeffrey. 1978. *Ngandi Grammar, Texts, and Dictionary*. Canberra: Australian Institute of Aboriginal Studies.
48. Heath, Jeffrey. 1984. *A functional grammar of Nunggubuyu*. Canberra: AIAS.
49. Merlan, Francesca. 1982. *Mangarayi*. London: Croom Helm.
50. Merlan, Francesca. 1994. *A Grammar of Wardaman: a Language of the Northern Territory of Australia*. (Mouton Grammar Library 11.) Berlin and New York: Mouton de Gruyter.

Pacific and Indonesia

51. Dixon, R. M. W. 1988. *A Grammar of Boumaa Fijian*. Chicago: University of Chicago Press.
52. Donohue, Mark. 1999. *A Grammar of Tukang Besi*. (Mouton Grammar Library 20.) Berlin and New York: Mouton de Gruyter.
53. Durie, Mark. 1985. *A Grammar of Acehnese on the Basis of a Dialect of North Aceh*. Dordrecht, The Netherlands, and Cinnaminson, NJ: Foris.
54. Harrison, Sheldon P. 1976. *Mokilese Reference Grammar*. Honolulu: University Press of Hawaii.
55. Josephs, Lewis S. 1975. *Palauan Reference Grammar*. Honolulu: University Press of Hawaii.
56. Lichtenberk, Frantisek. 1983. *A Grammar of Manam*. (Oceanic Linguistics Special Publication 18.) Honolulu: University of Hawaii Press.[5]
57. Mosel, Ulrike, and Even Hovdhaugen. 1992. *Samoan Reference Grammar*. Oslo: Scandinavian University Press.
58. Osumi, Midori. 1995. *Tinrin Grammar*. (Oceanic Linguistics Special Publications No. 25.) Honolulu: University of Hawaii Press.
59. Rehg, Kenneth L. 1981. *Ponapean Reference Grammar*. Honolulu: University Press of Hawaii.
60. Schachter, Paul, and Fe T. Otanes. 1972. *Tagalog Reference Grammar*. Berkeley and Los Angeles: University of California Press.
61. Sohn, Ho-min. 1975. *Woleaian Reference Grammar*. Honolulu: University Press of Hawaii.
62. Thieberger, Nicholas. 2007. *A Grammar of South Efate*. Honolulu: University of Hawaii Press.

Papua New Guinea

63. Davies, John. 1981. *Kobon*. Amsterdam: North-Holland.
64. Foley, William A. 1991. *The Yimas Language of New Guinea*. Stanford, CA: Stanford University Press.
65. Franklin, K. J. 1971. *A Grammar of Kewa*. Canberra: Linguistic Circle of Canberra.
66. Haiman, John. 1980. *Hua, a Papuan Language of the Eastern Highlands of New Guinea*. Amsterdam and Philadelphia: John Benjamins.

11.2.1.3 Other Resources

We also suggest reading any journal articles on particular problems in the language to be studied. And it would make little sense to set out on fieldwork without at

[5] Turner (1986) is a pedagogical grammar of Manam, in part derived from Lichtenberk's work, and in part based on additional fieldwork.

least skimming articles on related languages. There are also typological surveys that concentrate on a particular linguistic area, such as Abbi (2001:115–220), already mentioned. We recommend reading more general surveys first, and then moving to the more specific surveys. Lists of general surveys were provided in Section 4.3.1.

It is also useful to skim the typological questionnaires and other tools found on the website of the Max Planck Institute for Evolutionary Anthropology, Department of Linguistics: http://www/eva/mpg.de/lingua/tools-at-lingboard/tols/.php.

Finally, as a handy reference for linguistic terms, the following website is a good resource: http://www.sil.org/linguistics/GlossaryOfLinguisticTerms. Entries include short definitions, example constructions, and references.

11.2.2 Formal Marking Systems

Here we provide, in a nutshell, a description of a few formal morphosyntactic, morphological, or syntactic marking systems which a fieldworker must learn to recognize early on. We also provide references that we have found useful. Of course, there are other marking systems besides those we discuss here. Something as simple as constituent order is clearly a formal marking system in most languages, but this is easily distinguished by anyone, and need not be discussed here.

The term “marking” has been somewhat abused in linguistic terminology (see also Section 11.5). We all understand what is meant when we say that *dogs* is marked for plural, because *–s* is the plural marker, but is *dog* marked for singular? Semantically yes, but formally no. One could say that *dog* is not formally marked for number, but that it is interpreted as singular on semantic grounds. One could also say that the formal marker of singular is zero, but we suggest avoiding zero markers as much as possible. In our view, the morphosyntactic term “marking” should be limited to well circumscribed formal cases. Indeed, when using the term marking, it is always a good idea to define what is meant by it. The same remarks can be made about the term “coding”. For more on the difference between formal and functional marking, see Aikhenvald and Dixon (1998:60–61).

11.2.2.1 Head Versus Dependent Marking, or Locus

Before discussing this important typological parameter, let us define heads and dependents. We define the head of a word or construction as “the element that determines the syntactic function of the whole” word or construction (Payne 1997:31). What is not the head is the dependent.

This important typological parameter of head versus dependent marking was first described by Nichols (1986). “Locus” is the term proposed for this parameter in Bickel and Nichols (2007:193). Some languages mark the relationship between head and dependent on the dependent, and other languages mark this relationship on the head (Payne 1997:31; Whaley 1997:141–146). For example, in *John's dog*, *dog* is the head, *John* is the dependent, and the relationship between head and dependent – possession – is marked by the clitic *='s* on the dependent *John*. Thus English is dependent marking. But in the equivalent Western Apache phrase, *John bigósé*, the relationship between head and dependent is marked with the prefix *bi-* 'his/her' on the head *gósé* 'dog'. Thus Western Apache is head marking. Also, one should not assume that a language will be either head marking or dependent marking; some languages are both head and dependent marking.

The head versus dependent typology applies at different levels: phrasal, clausal, and sentential. Further discussion of this typology is in Section 11.4.5.

11.2.2.2 Inverse Marking

Inverse marking systems depend on a hierarchy of some type that can be reversed with inverse marking, so the effect of inverse marking depends on the hierarchy involved.

Some languages add an inverse marker to an inherently singular noun to make it plural, and add the same inverse marker to an inherently plural noun to make it singular. This is a case of number inverse marking.

In some languages, the inverse marker is added to the first person subject–second person object marker on the verb, so it comes to mean first person object–second person subject instead. This is an example of person and grammatical relation inverse marking. A discussion of languages with such inverse marking is in Klaiman (1992).

A syntactic inverse, sometimes called inverse voice, occurs in the Navajo language; the sequence: (Subject) NP (Object) NP Verb is interpreted as (Object) NP (Subject) NP Verb when the inverse is marked morphologically on the verb (Hale 1973). Whaley (1997:189) has an example of a syntactic inverse in Kutenai, another Native American language.

11.2.2.3 Switch Reference Marking

Switch reference is a device that marks subjects as being the same or different across more than one clause. It was first described for some Native American languages of the American West by Jacobsen (1967), then more broadly by Haiman and Munro (1983). The relationship between switch reference and topicality is discussed with New Guinean languages in Reesink (1983). Brief discussions with examples are in Foley and Van Valin (1984:339–354), Payne (1997:322–325), and Whaley (1997:276–279).

11.2.3 Lexical and Grammatical Categories

In the following two sections, we outline the major lexical and grammatical categories used in the morphosyntactic, morphological, and syntactic investigation of a language.

11.2.3.1 Lexical Categories

One of the earliest categorizing tasks that a fieldworker undertakes involves determining the lexical category of words, traditionally called parts of speech or form classes. In this section we provide a rough guide to lexical classification (based in part on Schachter 1985, and Schachter and Shopen 2007),[6] as well as some hints on how to identify lexical categories. For discussion of problems in identifying parts of speech, we recommend Cristofaro (2006:155–156), Enfield (2006), Munro (2007), and Vogel and Comrie (2000).

The following is a list of lexical categories, labeled as content or function words, and as open or closed classes. The fieldworker should be aware of the standard ways of distinguishing these lexical categories, keeping in mind that some of the criteria used are inconsistent. Croft (2001:63–83) presents interesting arguments against the existence of cross-linguistic lexical categories, in part based on such inconsistencies. However, we suggest that the descriptive fieldworker continue to use the familiar labels, while taking care to define them with precision. Avoiding lexical category labels altogether would be throwing the baby out with the bathwater.

Open class content words (exist in all languages)
1. Nouns
2. Verbs

Open or closed class content words (language dependent)
3. Modifiers
    3.1. Adjectives (modifying nouns)
    3.2. Adverbs (modifying verbs)
    3.3. Adverbs (modifying adjectives)
    3.4. Adverbs (modifying adverbs)
    3.5. Adverbs (modifying clauses or sentences)

Closed class function words
4. Noun adjuncts
    4.1. Adpositions
        4.1.1. Prepositions
        4.1.2. Postpositions
        4.1.3. Circumpositions
    4.2. Case "particles" (i.e., not morphological affixes)
5. Verb adjuncts
6. Conjunctions
    6.1. Coordinating conjunctions
    6.2. Subordinating conjunctions
        6.2.1. Complementizers
        6.2.2. Relativizers

[6] Payne (1997:32) departs from the norm in calling lexical categories "grammatical categories".

        6.2.3. Adverbializers
7. Pro-forms
    7.1. Pronouns
    7.2. Pro-verbs
    7.3. Pro-adjectives
    7.4. Pro-adverbs
    7.5. Pro-predicates
    7.6. Pro-clauses
    7.7. Pro-sentences
8. Other, or residue

Following the American structuralist tradition, it is often useful to distinguish open lexical categories by morphological criteria. For example, nouns can be distinguished from verbs (and perhaps from pronouns, adjectives, and adverbs as well) because of characteristic inflectional morphology, such as plural marking or case marking. Syntactic criteria, such as the ability to co-occur with another lexical category, can also be used. Closed classes are typically distinguished by syntactic criteria only. Some lexical subcategories, such as interjections and ideophones (usually a subcategory of verb or adverb (Hinton et al. 1994; Jendraschek 2001)) can often be distinguished by phonological criteria as well.

Categories (3–8) are not universally present, though one can generally identify several types of function words in addition to the open classes. But the question of which type of function words might be expected to exist is not easy to answer cross-linguistically. The lesson for fieldworkers is to define classes of function words by criteria that are very explicit.

Closed classes, with the exception of Pro-forms, are typically not inflected. The inflected prepositions of Celtic are an exception; these are the result of the historical combination of a preposition with clitic pronouns (Anderson 1985b:201). The inflected conjunctions of Dutch dialects (Haegeman 1992) are another unusual case; here, the conjunctions agree in number and person with the following verb.

The fieldworker needs to be aware that morphological criteria and syntactic criteria may not necessarily coincide; for example, some English adjectives fit the morphological frame for adverbs (occurring with the adverb-deriving *-ly* suffix), but fit the syntactic frame for adjectives (occurring as nominal modifiers, as in *lowly manger*).

There is obviously a wealth of literature discussing each of these categories in typological and in theory-specific terms. Our goal in the following paragraphs is to provide the fieldworker with additional information that we have found useful in data gathering and analysis. We also provide references to sources where the reader can find out more on each category.

Nouns and Verbs

It seems that all languages make some sort of distinction between nouns and verbs. This is not a morphosyntactic requirement; rather, it is related to the fact that human language needs to distinguish, in terms of information structure, between a "topic" and a "comment", or in terms of predicate calculus, between "arguments" and "predicates"

(Payne 1997:174), or semantically, between "reference", and "predication". Dixon (2010b:37–61) makes a strong but somewhat belabored case supporting such a distinction between noun and verb classes.

Needless to say, isolating languages – i.e., languages with very little morphology – will not make a morphological distinction between nouns and verbs. Some languages – such as the languages of the Salish and Wakashan language families of the North American northwest coast (Jacobsen 1979; Kinkade 1983; Kuipers 1968) –, have quite a bit of morphology, but still make very little morphological distinction between nouns and verbs. In such a case, the distinction is made on the basis of syntactic evidence. Finally, some morphologically complex languages tend not to make a morphological distinction between nouns and stative verbs, but do make a morphological distinction between nouns/stative verbs and active verbs. Languages of the Siouan family are an example.

Nouns

Nouns can often be characterized as Proper versus Common, Possessable versus Nonpossessable, Count versus Mass, and Human versus Nonhuman. A common subtype of noun is the Relational noun which occurs in constructions that function like adpositions, for example, English *front* and *back*, used in *in front of* and *in back of* (Bickford 1998:108).

Verbs

The defining characteristics of verbs are captured by the semantic classification of Vendler (1967), which has stood the test of time. On the basis of their semantic, syntactic, and aspectual behavior, Vendler distinguished verbs as referring to States, Activities, Accomplishments, or Achievements. Exemplification and further discussion of these classes is given in Dowty (1979), Foley and Van Valin (1984:36–47), and Payne (1997:55–61). The most detailed syntactic classification of verbs that we know of in any language is Levin (1993), for English.

Adjectives and Adverbs

In many languages, adjectives are a subcategory of either verbs or nouns. Additionally, in many languages, adverbs are derivable from verbs or from adjectives. Finally, in some languages, the distinction between adverbs and adjectives is not clear-cut; the morphological distinction between adjectives and other modifiers that is so clear in Indo-European languages is not a universal.

One defining characteristic of adjectives and adverbs is their ability to occur in comparative and superlative constructions (Payne 1997:88), assuming these constructions exist in a particular language.

Adjectives

Adjectives, as closed or open classes, are surveyed in Dixon and Aikhenvald (2004). They suggest, in this volume, that some sort of adjective class is a linguistic universal on par with nouns and verbs, and it does, indeed, seem to be the case that one can always find some morphosyntactic criteria by which one can distinguish some adjective-like category, however subtle. Opinions remain divided as to whether adjectives are open or large and important classes in all languages. Dixon (2010b:62–114) reiterates arguments for distinguishing an Adjective class. Other discussion of adjectives is in Lindsey and Scancarelli (1985).

Adverbs

Adverbs can be subclassified according to what they modify (Verbs, Adjectives, or other Adverbs). They can also be classified semantically into Adverbs of Degree, Frequency, Manner (of which Ideophone can be a subtype), Place or Location, Directional, and Time Adverb.

Sentence-level or clause-level Adverbs tend to be Modal (Epistemic or Deontic), Mirative, or Evidential.

Some languages mark a formal distinction between the different subcategories of adverbs, but most do not. Descriptive traditions often have distinct names for adverbs derived from verbs. In the Indo-Europeanist tradition, they are called Gerunds or Adverbial Participles; the term Converb is used for Altaic and other languages of the Caucasus and Central Asia (Haspelmath 1995:2–5). However, the terms Gerund, Participle, or Converb should not be used to refer to inflected forms of verbs; if they are inflected, then of course they should not be considered a type of adverb at all, since they belong to the lexical category of Verbs.

Verbal "particles", like those that occur in English (e.g., *up* or *in* in *sign up to a website*, or *sign in to a website*) and other Germanic languages are a special type of Adverb which modifies Verbs. However, any sentence adverb that is short and difficult to describe semantically is likely to be called a particle or a discourse marker. Useful surveys of particles in German (generally called "modal", though not all their functions are modal) are in Abraham (1991a, b).

Adverbial constructions are surveyed in van der Auwera and O'Baoill (1998).

Noun Adjuncts

This term covers "Adpositions", as this is a well-accepted cover term for Prepositions and Postpositions. One also expects the existence of Circumpositions, i.e. an "adposition" composed of a pre-posed as well as a post-posed element. In many languages, case marking elements or semantic role marking elements are formally Adpositions. The distinction between Adpositions and Adverbs is not always clear.

Verb Adjuncts

This category includes auxiliary verbs, also called "helping verbs", which can have their own inflectional characteristics and can co-occur with other verbs. English has a particularly unusual system of auxiliaries with "Do-support". Descriptive traditions may have other names for auxiliary verbs; for example, in South Asian languages, the terms "explicator verb" and "vector verb" are used (Abbi 2001:45). Other terms are "light verb", "function verb", "secondary verb", "combined verb", and "aspectual verb" (Sen-Gupta 1978; Schultze-Berndt 2006:361). Surveys of auxiliary verb constructions are in Heine (1993) and Anderson (2005). A discussion from the point of view of grammar writing is Schultze-Berndt (2006).

Conjunctions

In many languages, Conjunctions are not clearly distinguished, formally, from Adverbs or from Adpositions. Furthermore, the distinction between Coordinating and Subordinating Conjunctions is often tenuous. In general, subordinating conjunctions can embed clauses to create complement, relative, or adverbial clauses. Coordinating conjunctions, of course, join like constituents.

Pro-forms

There are several different types of pro-forms, of which Pronouns are the most recognizable. Pronouns are of the following types: personal, reflexive, reciprocal, demonstrative, relative, indefinite, and interrogative. We will return below to the special status of indefinite, demonstrative, and interrogative pronouns. An interesting characteristic of pronouns is that they often retain gender and case inflection where it was lost or never existed in the rest of the system. Pronouns are discussed in Dixon (2010b:189–223).

Pro-verbs (not to be confused with proverbs) typically are interrogative verbs meaning things such as 'to do what?', 'to say what?', 'to be who?', or 'to be how?'.

Pro-adjectives may be found in expressions that are translatable as 'a <u>what</u> dog?', to which the answer would be 'a <u>black</u> dog', or the like.

Pro-adverbs are typically interrogative adverbs, such as *when, how, where,* and *why*. The answer would typically contain an adverb or adverbial construction.

Pro-predicates, Pro-clauses (such as question tags), and Pro-sentences (such as interjections) are discussed in Schachter (1985), and Schachter and Shopen (2007).

Words That Cut Across Lexical Categories

Numerals, non-numeral quantifiers, determiners, indefinites, and interrogatives are not easy to classify, as they do not fit neatly into any one specific lexical category. What they seem to have in common is that they cut across our list of lexical categories, and they provide information that is beyond the domain of syntax, such as logical operator information in the case of numerals and quantifiers, and pragmatic information in the case of determiners, articles, demonstratives or deictics, interrogatives, and indefinites.

Consider, for example, numerals. Cardinal numerals (*one, two, three*, etc.) can function as Pronouns or Modifiers; Ordinal numerals (*first, second, third*, etc.) are typically Modifiers but can also function as Nouns and Adverbs; Aggregative numerals (*a group of two, a group of three*, etc.) are typically Nouns (Abbi 2001:134); Multiplicative numerals (*once, twice, three times* or *thrice, four times*) are Adverbs of Frequency; and Distributional numerals (*one by one, two by two, three by three; in ones, in twos, in threes*) are Adverbs. Numerals and nonnumeral quantifiers can also be partitive and nonpartitive (Abbi 2001:134).

Quantifiers and Determiners can also be regarded as cutting across lexical categories. They can be considered as Modifiers, but they do not fit unequivocally into this category. Furthermore, they can exhibit some confusing form-function overlaps: for example, some non-numeral quantifiers can function as Determiners. Deictic Determiners are also called Demonstratives; any non-deictic determiner is often called an Article (Payne 1997:102). However, how exactly determiners differ from modifiers (except in obvious ways such as position in the Noun Phrase in English) is not often clear. Demonstratives are surveyed in Diessel (1999), and in Dixon (2003, 2010b:223–261).

Constituents that convey pragmatic distinctions – such as speech-act (declarative vs interrogative mood), definiteness, or deixis – also typically cut across lexical categories. So we can talk about interrogative, demonstrative, and indefinite Adjectives, Adverbs, or Pro-forms. Thus interrogatives and question words are not independent lexical categories; rather, they can be Adjectives (*which man?*, *what color?*), Pro-Adverbs (*how?, why?, where?, when?*), Pronouns (*what?, who?*), and in languages other than English, Pro-verbs, to be translated as in 'to be called what?', 'to say what?', 'to be how?', or 'to do what?'.

The "Other" or Residue Category

Finally, if the fieldworker comes across a unique type of word which s/he cannot fit into the above classification, it is not a mortal sin to put it into the "anything that is not a noun-verb-modifier-adjunct-conjunction-or-proform" category. Examples of things that might have to be put in a "residue" category" (assuming none of these are actually morphological) are:

- A linker, connecting a Noun and its Modifier (assuming it is not a Conjunction or Adposition)
- A copula, connecting a subject and a predicate nominal or adjectival (assuming it is not a Verb or Pro-form)
- A predicator, turning a nominal or adjectival form into a predicate, without the need for a subject (assuming it is not a Verb or Pro-form)
- An existential or presentative particle (assuming it is not a Verb or Pro-form)
- A relative clause marker (assuming it is not a Conjunction or Pro-form)
- A set of "Modal particles", perhaps mood or discourse markers (Abraham 1991a, b; Wierzbicka 1976, 1986b) (assuming they are not Adverbs)

- A negator (or a prohibitive, if it is imperative or optative) (assuming it is not an Adverb)
- A politeness marker (assuming it is not a Verb, Modifier, or Proform)

It is easy to abuse this category, so it is best to limit the number of elements are classified this way. As argued in Section 11.5, one should as much as possible avoid positing a category of "particles", a wastebasket category for uninflected words.

11.2.3.2 Grammatical Categories

An equally important task that a fieldworker undertakes involves determining the grammatical categories at work in the morphosyntax, morphology, and syntax of a language. Many speakers have an intuitive notion of what a grammatical category is in their language. An English speaker will easily recognize number (singular vs plural); a French or Spanish speaker will easily recognize gender (masculine vs feminine). However, grammatical categories are sometimes not very obvious, and it can be difficult for a fieldworker to determine which ones are relevant in a particular language. In this section, we provide a rough guide to grammatical categories, and some hints on how to identify them.

Grammatical categories are different for noun phrases (Section "Grammatical Categories Characteristic of Noun Phrases"), for verb phrases (Section "Grammatical Categories Characteristic of Verb Phrases"), and for clauses or sentences (Section "Grammatical Categories Characteristic of Clauses or Sentences").

Grammatical Categories Characteristic of Noun Phrases

The fieldworker should be prepared to look for the following characteristics of nouns or noun phrases:

- Semantic Role
- Pragmatically Marked Status
- Case
- Possession
- Number
- Noun class or gender
- Noun classification
- Deixis

These are discussed in detail below.

Semantic Roles

Semantic roles (or theta-roles, as they are called in the generative literature) were first posited by Gruber (1965) and Fillmore (1968), and are, technically speaking, a semantic distinction rather than a morphosyntactic one. They are widely discussed in the typological literature, as in Foley and Van Valin (1984:28–36), Andrews (1985, 2007a), Payne (1997:49), Whaley (1997:64–67), and Bickford (1998:107). The following distinctions are useful:

- Agent and patient roles, which correspond syntactically with arguments
- Participatory roles, which correspond syntactically with complements, and which can be further subclassified into Directional roles, which include Goal and Source, and Nondirectional roles, which include Inner locative i.e., location of a participant, Experiencer, Recipient, Theme, Cause, and Instrument
- Circumstantial roles, which correspond syntactically with adjuncts, and include Benefactive, Outer locative (the place where some action or event occurs), Reason, Comitative, and Temporal

No language attempts to code all possible semantic roles, since the number of proposed semantic roles is nearly infinite. However, in all languages, we see that morphological or syntactic operations are sensitive to semantic roles to some extent, so it is important for the fieldworker to be aware of the range of semantic roles and how they can be encoded.

Coding strategies for semantic roles include: constituent order; affixation or cliticization (often simply called case-marking); and cross-referencing or agreement, with possibilities including agreement in person, number, case, noun class, or gender; or some combination of these. See Andrews (1985, 2007a) for further discussion.

Pragmatically Marked Status

Pragmatically marked status is technically a pragmatic distinction rather than a morphosyntactic one. It is discussed in Li (1976), Foley and Van Valin (1985:287), Andrews (1985, 2007a), Payne (1997:282), Kiss (1998), and Lyons (1999). The following distinctions are relevant to Pragmatically marked status:

- Presupposed versus Not presupposed.
- Focus versus Not in focus. Note that negation is inherently a type of focus.
- Definite or Identifiable versus Indefinite or Not identifiable (Givón 1978).
- Given versus New (Prince 1981).
- Topic or Theme versus Comment or Rheme.
- Generic versus Specific.
- Referential versus Nonreferential (Givón 1978).

Particular care needs to be taken to distinguish between two or more pragmatic oppositions which are indicated by the same morphosyntactic encoding. To correctly understand the distribution of morphosyntactic elements, the fieldworker should take care to keep the following pairs of oppositions apart:

- Presupposed/Not presupposed versus Informational Focus/Not in focus
- Presupposed/Not presupposed versus Definite/Indefinite

- Given/New versus Topic/Comment
- Given/New versus Definite/Indefinite
- Generic/Specific versus Definite/Indefinite
- Referential/Nonreferential versus Definite/Indefinite
- Generic/Specific versus Referential/Nonreferential

Morphosyntactic coding strategies for such pragmatically marked statuses include: articles, particles, or clitic-like elements, often called discourse markers; syntactically marked constructions such as cleft and pseudocleft (Payne 1997:271; Foley and Van Valin 1985:358–363); passive voice; and constituent order changes such as left dislocation, right dislocation, extraposition, preposing, or postposing (Foley and Van Valin 1985:355–358; Foley 2007; Payne 1997:273). Left dislocation and preposing are particularly common strategies for marking Presupposed information, Given information, and Topic; however, they are not used to mark Focus.

The pragmatically marked status of a constituent can, of course, also be indicated by non-morphosyntactic means, such as special stress or intonation.

Case

Case is a formal device used to mark grammatical relations such as subject and object, as well as various complements and adjuncts. Morphosyntactic strategies for marking case can be morphological (inflectional), by clitics (see Section 11.3.3.1), by use of a separate word (often called a particle), or syntactic i.e. through constituent order. Case typically overlaps with semantic role marking and interacts in complex ways with pragmatic marking, so it may be difficult to distinguish a potential case marking system from a semantic role or pragmatic system (Chelliah 2009; Gaby 2008). Common case distinctions and labels are: nominative/accusative; ergative/absolutive; genitive; dative; instrumental; comitative; equative; vocative; and various adverbial cases such as locative, allative, illative, elative, ablative, adessive, inessive, and abessive. A useful list of case label definitions can be found online at http://en.wikipedia.org/wiki/List_of_grammatical_cases.

Ergative/absolutive or nominative/accusative-like "patterning" in some part of the syntax or lexicon should not be confused with morphological ergative/absolutive or nominative/accusative case marking. Dixon (2010a:78) warns about the misuse of the term "ergative".

Basic references on case are Blake (1994), Payne (1997:100), Whaley (1997:166–168), and Malchukov and Spencer (2008). For the relatively uncommon phenomenon of double case, Plank (1995) is the best collection of studies. Campe (1994) is a bibliography on case and semantic roles.

Possession

There is wide cross-linguistic variation in terms of what can and what cannot be possessed, so the fieldworker should be familiar with the encoding of possession in related languages and in the cultural and geographic area of the language being investigated. Also, alienable possession is often formally distinguished from inalienable possession. See Ultan (1978b), Payne (1997:104–105), and Bickford (1998:73) for further discussion.

For an expression that translates into English as *John's dog*, morphosyntactic coding strategies for possession can include: simple juxtaposition of the possessor and possessum (or possessee) ("John dog" or "dog John"); marking of the possessum ("John his-dog"); marking of the possessor (as with a genitive case "John's dog" or "dog John's", a preposition "dog of John", or a postposition "John of dog"); or marking of both the possessor and the possessum ("John's his-dog"). Thus, as for case marking, possession can be marked morphosyntactically through affixation, cliticization, with a separate word, or through constituent order.

Possession is surveyed in Baron et al. (2001), and in Dixon (2010b:262–312).

Number

Number is discussed in Corbett (2000) and Payne (1997:96–100). The basic distinctions are well known: singular, plural, dual (two), and paucal (a few). Trial (three), as distinguished from paucal, is attested only rarely. A very large proportion of the languages of the world do not mark number as an obligatory category, and instead indicate number, if needed, with numerals, non-numeral quantifiers, through Noun classification, or by marking on the verb. Some languages have a singulative which isolates an element from a typical pair or group. In such a language the term "eye" would mean 'a pair of eyes' and "eye-SINGULATIVE" would mean '(one) eye". Other languages distinguish a collective plural (referring to a group of items together) from a distributive plural (referring to a plurality of items separated from each other in space or time). Inverse number marking was discussed in Section 11.2.2.2.

Coding strategies for marking number can be morphological, through cliticization, or through use of a separate word.

Noun Class or Gender

Noun class or Gender is discussed in Dixon (1986), Corbett (1989, 1991, 2007), and Payne (1997:105). An important characteristic of Noun Class or Gender is that, synchronically, it is not semantically motivated. This is what distinguishes Noun Class or Gender from Noun Classification, to be treated next. Corbett (1991:7–32) discusses strict semantic systems and predominantly semantic systems as the basis for assigning gender. However, even so-called strict semantic systems require at least some knowledge of the extra-linguistic cultural context in which the language is spoken, and such knowledge is not semantic.

Two types of gender – sex gender (masculine/feminine, masculine/feminine/neuter, or common/neuter) and animacy (animate/inanimate) – are considered to be semantically motivated systems because they derive from diachronically motivated distinctions, but even these exhibit synchronic exceptions. A system with many genders is called a noun class system; these tend to be less semantically non-motivated than animacy or sex gender. Examples are the noun classes in Bantu and other African language families. It is possible for one language to have gender in one part of its system, and noun class in another (Aikhenvald and Dixon 1998:59).

Noun Class or Gender can be marked morphologically, through cliticization, or with a separate word. The most common strategy is affixation.

Noun Classification

Noun classification should not be confused with Noun Class or Gender. Aikhenvald and Dixon (1998:58) use the term Reference Classification as a cover term for Gender, Noun Class, and Noun Classification. However, we believe that using such a cover term might obscure the crucial differences between Noun class or Gender on the one hand, and Noun Classification on the other.

Noun classification is discussed in Dixon (1986), Craig (1986), and Aikhenvald (2000, 2004a). Noun classification is semantically motivated, synchronically and without exception, according to observable characteristics of animacy, humanness, sex, tangibility, size, value, shape, number, consistency or compactness, dimensions on the horizontal axis and/or vertical axis, and honorific status. Noun classification can overlap with a number or honorific system. The fieldworker needs to be aware that noun classification according to sex or animacy is systematically semantically motivated, whereas noun class and gender based on sex or animacy are not systematically semantically motivated. It is also possible for a language to have both Noun Class or Gender, and Noun Classification (Aikhenvald and Dixon 1998:59).

Morphosyntactic coding strategies for noun classification are marking on the noun itself, morphologically or through cliticization, or by a special adjective, numeral or quantifier, generally simply called a classifier, or by the morphology or stem shape of the verb of which the noun is an argument.

Deixis

Deixis, technically speaking, is a pragmatic rather than a grammatical category. Useful discussions of Deixis can be found in Levinson (1983:54–96), Anderson and Keenan (1985), and Hanks (2005). Distinctions relevant to deixis generally interact with the category of Person. Third person can be further distinguished as proximate/obviative, or even proximate/obviative/further obviative (Anderson and Keenan 1985:262). See Section "Grammatical Categories Characteristic of Verb Phrases", under "Subject Person and Number" for further discussion of Person. Deixis also interacts with number, gender, and the social rank of the participants. One can distinguish spatial deixis and temporal deixis. Spatial deictic systems range from very simple systems like English: *this/that, here/there* (but note the archaic system *this/that/yon, here/there/yonder*) to the very complicated demonstrative systems of Eskimo and many other Native American languages (Boas 1911:40–41; Woodbury 1993). Temporal deictic systems commonly make a three-way distinction, e.g., *tomorrow/today/yesterday*, but there are also many languages that have only one word for both 'tomorrow' and 'yesterday'; that is, a system contrasting 'today' with 'a day away from today', either plus or minus.

Spatial deixis is further surveyed and studied in Levinson (2003), and Levinson and Wilkins (2006).

It is also useful to distinguish between anaphoric deixis, in which the deictic element (in bold) points backwards, as in … *and John said* ***the same thing***, from cataphoric deixis, in which the deictic element (in bold) points forwards, as in *The winners are* ***the following****: …*.

Another useful distinction is that between absolute deixis versus relativized deixis. Relativized deixis typically occurs in indirect speech. An example of absolute temporal deixis is *He was sick*, where the past tense of *was* refers to a time prior to the moment of speech. But in *John will say he was sick*, the past tense reference of *was* is not absolute, but relative to the event *John will say* (Anderson and Keenan 1985:301). The issue of relativized deixis is further discussed in Anderson and Keenan (1985:301–307).

Deixis can be encoded morphologically, through cliticization, or syntactically by the use of Modifiers and Pro-forms.

Grammatical Categories Characteristic of Verb Phrases

The following is a list of the grammatical categories that are characteristically indicated on verbs or verb phrases:

- Tense
- Aspect
- Mood/Mode/Modality
- Voice
- Valence
- Polarity
- Control/Noncontrol
- Subject person and number
- Object person and number
- Associated motion

All of the above can be encoded morphologically, though cliticization, or syntactically by the use of separate words (Adjuncts such as auxiliaries, or Pro-forms) or through constituent order.

Tense

The categories Tense, Aspect, and Mood/Mode/Modality are often intimately connected, and are often discussed together as Tense/Aspect/Mood (TAM) categories. For example, there is considerable overlap of irrealis mood with future tense, and realis mood with past tense. Comrie (1985a), Chung and Timberlake (1985), Payne (1997:233), Whaley (1997:204–210), and Timberlake (2007:304–315) are surveys of Tense. A few languages, such as Hupa, a Native American language of California, have tense marking on nouns, but tense marking on nouns is never as systematic or elaborate as tense marking on verbs.

As discussed above for Deixis, in many languages tense can be absolute or relative. The basic absolute tense distinctions are present/past/future. Some languages make distinctions between several different past tenses: remote, regular, and immediate (and in a few cases even subtler distinctions), or between several different future tenses: remote, regular, and immediate. The basic relative tense distinctions (in the prevailing Indo-Europeanist terminology) are: perfect, i.e., a past action with respect to the present; pluperfect, i.e., a prior action with respect to the past; and future perfect, i.e., a prior action with respect to the future (Nida 1949:167). Since a past action with respect to the present generally means that the action continued up to the present, it is clear that the English perfect has some aspectual connotations.

Aspect

Aspect is discussed and surveyed in Nida (1949:168), Comrie (1976), Dahl (1985), Chung and Timberlake (1985:218), Payne (1997:233), Whaley (1997:210–216), and Timberlake (2007:284–304). In general, cross-linguistically, aspectual distinctions are found more often than tense distinctions, and they also tend to be more complicated than tense distinctions. Also, it is not always easy to tease apart tense and aspect. Types of aspectual distinctions include:

- Imperfective versus Perfective or Incompletive versus Completive (i.e., the action or process described by the verb is viewed as not completed/completed), the most common distinction cross-linguistically
- Momentaneous versus Progressive or Continuous or Continuative versus Prolongative (a potential three-way distinction: the action or process goes on for a very short time/goes on for some period of time/goes on for a long time)
- Punctual or Punctiliar versus Durative (i.e., the action or process has no duration/has duration)
- Inchoative or Inceptive versus Cessative or Terminative (i.e., the action or process is beginning/is ending); these distinctions are sometimes called phasal
- Iterative versus Semelfactive (i.e., the action or process occurs several times/one time); subtypes of Iteratives include the following:
    - Repetitive (i.e., the action or process is repeated at regular intervals)
    - Frequentative (i.e., the action or process is repeated frequently)
    - Habitual, Habituative, Customary, or Usitative (i.e., the action or process is carried out as a habit or custom)
    - Semeliterative (i.e., the action or process happens one more time)
    - Seriative (i.e., separate actions or processes are carried out in a series, one after another)

The German term Aktionsart is often used to talk about the lexical or inherent aspectual characteristics of a verb. Thus one can say that the verb 'to read' is inherently durative or that it has a durative Aktionsart, and that the verb 'to sneeze' is inherently punctual or that it has a punctual Aktionsart (Whaley 1997:214–215). We consider Aktionsart to be a semantic characteristic of particular verbs, and therefore do not regard it as a grammatical category. The fieldworker should bear in mind that much descriptive confusion can occur when a grammar writer does not clearly distinguish aspect from Aktionsart. A recent treatment of the theory of aspect is Sasse (2002).

The fieldworker will also find it useful to consult the grammatical tradition relevant to the language under study for the accepted terminology for tense and aspect (Foster 1986). However, as much as possible, it is preferable to give standard equivalents.

Mood/Mode/Modality

Mood/mode/modality is discussed and surveyed in Chung and Timberlake (1985:241–257), Palmer (1986), Payne (1997:233), Whaley (1997:219–226), and Timberlake (2007:315–330). For some scholars, mode, mood, and modality all mean the same thing; for others there are differences between them. We will use Mood as a cover term for all three. Within mood, one can distinguish Epistemic and Deontic moods. Within the verb or verb phrase, Epistemic mood is primarily the realis/irrealis distinction. The subjunctive and conditional moods are subtypes of irrealis. Within Deontic mood, we can distinguish: obligatory, permissive, volitional, desiderative or optative, and prospective or intentive moods (Nida 1949:169).

Voice

Voice should not be confused with valence or valency (see next paragraph). Valence has to do with the number of arguments a verb requires. Voice has to do with the functional relationship between the arguments of a verb. Valence can be inherent to a particular verb, while Voice is not.

The best known voice distinction is active/passive, the passive being the marked element of the pair (Keenan 1985a:249; Keenan and Dryer 2007:325). Some languages simply do not have passives. Many languages only have agentless passives, and if the agent is specified (e.g., in the English *by*-phrase) agentive passives are often much less frequently used than agentless ones. Foley and van Valin (1985:299–335) discuss the function of the passive, which can be to either foreground or background different arguments.

Another voice distinction is active/antipassive, which is common in ergative languages. Foley and van Valin (1985:335–347) discuss the functions of the antipassive, which can be again be either foregrounding or backgrounding.

Other types of voice (but see also Valence below) are the reflexive (acting upon oneself) (Geniušienė 1987; König and Gast 2008); the reciprocal (acting upon each other) (König and Gast 2008); the middle (acting for or on a part of oneself) (Kemmer 1993); the causative (causing someone else to act); and the benefactive (acting for the benefit of another). Malefactives (acting to the detriment of another) are much less common than benefactives.

For useful discussions of voice, see Foley and Van Valin (1985), Keenan (1985a), Klaiman (1991), Bickford (1998:313), and Keenan and Dryer (2007). A very systematic account of possible voice distinctions is Mel'čuk (2006). While we agree that the term antipassive is not a revealing one, we are not convinced by Mel'cuk's (2006:191–193) argument that the antipassive is not a voice, since like a passive, an antipassive can have foregrounding or backgrounding functions.[7]

Valence or Valency

Valence (or valency) should not be confused with Voice. However, as seen below, valence-changing devices overlap with Voice to a large extent. Valence is discussed in Payne (1997:169–222) and in Whaley (1997:183–200).

The Valence of a verb is the inherent capacity of the verb stem to take a certain number of arguments. These arguments can generally be described in terms of Subject or Object, but valence is also relevant in languages where Subject and Object are not demonstrable grammatical categories.

The basic valence distinctions are intransitive (no direct object)/transitive (one direct object)/ditransitive, also called bitransitive (one direct object and one indirect object). For most languages, valence only involves the object or objects of the verb. In some languages versus, some verbs cannot take subjects (e.g., impersonal verbs in Lakota, a Siouan language), and in this case the obligatory absence or presence of a subject can also be a matter of valence.

Ambitransitive or labile verbs are verbs which can be either intransitive or transitive. The subtypes of ambitransitive verbs are discussed in Dixon (2010a:77–78, b:123–126, 143–147).

It is theoretically possible to distinguish between the Valence of a verb – i.e., the inherent capacity of the verb stem to take a certain number of arguments – and the Transitivity of the verb, i.e., how many arguments it has after valence-decreasing or valence-increasing processes have been applied to the stem (discussed below). In some cases, this distinction can be useful. The issue of transitivity is surveyed in Dixon (2010b:115–158).

In the case of ditransitive verbs, it is not always easy to decide which argument is the direct object and which is the indirect object. It can be useful to talk about ditransitive verbs with two direct objects, with additional syntactic testing required to determine the relative syntactic status of the two objects. In Latin, for example, some verbs of asking or teaching can take two objects marked with the accusative

[7] We also disagree with Mel'cuk's (2006:174–175) proposal that the term "ergative construction" should mean "the predicative construction in which the Subject is marked by a case other than the nominative". Just because such a sloppy definition is common does not mean it should be accepted.

case: *mē sententiam rogāvit* 'he asked me my opinion', *docēre puerōs elementa* 'to teach children their ABC's' (Allen and Greenough 1931:246).

Valence becomes an interesting avenue of investigation when the language has devices which can increase, decrease, or rearrange it. Valence changing processes are surveyed in Dixon and Aikhenvald (2000).

Valence increasing devices are causatives, applicatives, and possessor-raising or ascension (Payne 1997:172). Functionally, some valence increasing devices are not distinguishable from alternations in Voice.

Valence decreasing devices are reflexives and reciprocals (for obvious semantic reasons), middles, subject or object omission, passives, antipassives, object demotion, and subject or object incorporation (Payne 1997:172). Here again, functionally, some valence increasing devices are not distinguishable from alternations in Voice.

Valence rearranging devices, called valence transposing devices by Whaley (1997:189), are processes such as the syntactic inverse (discussed in Section 11.2.2) and dative shift (Foley and Van Valin 1985:347–354, Payne 1997:192, Whaley 1997:189–90).

In their discussion of valence rearranging devices, Foley and van Valin (1985: 291–299) also mention converse predications, although these involve lexical pairs and do not involve any morphosyntactic device. But, like actual valence transposing devices, they certainly have semantic and pragmatic consequences. An example of a converse predication is the construction pair: *X bought Y from Z* versus *Z sold Y to X* (Foley and Van Valin 1985:291). There are languages, like Western Apache, where such pairs are rare, and both "sell" and "buy" are derived from the same verb stem.

Polarity

Polarity refers to the affirmative/negative distinction. In all languages the unmarked situation is, of course, affirmative polarity, with negation being the marked option. Negation can be marked through derivational or inflectional morphology or through a separate word. Word negation, for example, is often marked derivationally or lexically – in antonyms, for example – but can also be marked through separate words or clitics. The scope of negation and negative polarity items like *anymore* in English need to be considered as part of the syntactic analysis of a language. A good survey of negation is Payne (1985a).

Control/Noncontrol

The parameter of control or lack of control over an action is another possible morphosyntactic grammatical category which can be coded by inflectional, derivational, or lexical means. This is the parameter that lexically distinguishes verbs like *see* from *look* and *watch*; *hear* from *listen*; *pour* from *spill*; *tell* from *blurt out*. These distinctions are typically lexical in Indo-European languages, but can be derivational or even inflectional in other language families. For example, the parameter control/noncontrol is pervasive in the North American Salish family and is not only lexical, but occurs in many derivational or inflectional affixes as well (Thompson

1979, 1985). However, there is no particular affix that by itself means 'control' or 'noncontrol' (Timothy Montler, p.c.). It is important not to confuse this distinction with the active/stative distinction, or with a lexical "carefully, gently versus carelessly, roughly" distinction, which exists in the Athabascan languages (Rice 1989:784–87).

Subject Person and Number

The person and number of the Subject of the verb is a very common inflectional distinction. Number can be marked lexically, i.e., by suppletive verb stems that indicate subject number. Subject person can be marked by suppletion as well, as in English, *(I) am, (you) are, (s/he) is*, but this phenomenon is cross-linguistically less common than number suppletion. Subject person distinctions are of course first, second, and third person, as well as duals and plurals (and more rarely paucals or trials) of these person categories. Strictly speaking, there is no such thing as a true first person plural, since there cannot be more than one speaker, although it is possible to report on an event involving first person plural referents. Many languages distinguish between a first person inclusive and a first person exclusive (Haas 1969). The first person inclusive refers to 'we (the group the speaker speaks for or belongs to)' or 'I (the speaker) and you (the addressee or addressees)', (which, in some languages, when involving one addressee and one speaker, can be thought of as a first + second person). The first person exclusive refers to 'we (the group the speaker speaks for or belongs to) but not you (the addressee or addressees)'. Both inclusive and exclusive first persons can thus be dual or plural in number, but only a first person exclusive dual entails a dual subject: 'we (a group of two the speaker speaks for or belongs to)'.

The most recent survey of Person is Siewierska (2004), and a summary is Payne (1997:44–45). It is always the case that first and second persons are phonologically and syntactically distinguishable from the third person. It is very common cross-linguistically for number distinctions in the third person to not correspond with number distinctions in the first or second persons. Standard English is fairly unusual in having a number distinction everywhere in the personal pronoun system except in the second person.

Many languages have pragmatic constraints against personal pronouns, and instead prefer titles or kinship terms to refer to first, second or third persons, and make heavy use of inference. The best studied examples are Japanese and Thai, but quite a few other Southeast Asian and East Asian languages display the same sort of pronominal reference avoidance strategies (Foley and Van Valin 1984:324).

It is also not uncommon for languages to distinguish between a primary third person, and another third person who may have to be distinguished through politeness or taboo. This second third person is sometimes called a fourth person.

Another possibility is that one third person can be marked proximate and the other third person obviative.

Most accounts do not distinguish between the concepts of obviative and fourth person. It is often necessary to distinguish them, as follows. In a system with a third/fourth person distinction, there is no strict requirement that, of two third persons, one has to be marked third and the other one fourth. Furthermore, if a fourth person is selected, it tends to refer to the same person throughout the discourse. In a proximate/obviative system, one third person must be obligatorily marked as proximate, another third must be obligatorily marked obviative, and decisions as to which third person is proximate and which third person is obviative can change in the same stretch of discourse.

As a result, in languages with a third/fourth person distinction, the selection of one person as fourth is typically based on cultural constraints, such as taboo or politeness. In languages with a proximate/obviative third person system, constraints on which third person is proximate and which obviative are pragmatic, having to do with first reference and topicality, and there are pragmatic rules for reversing the assignment of proximate and obviative. Some languages can distinguish between an obviative and a further obviative (Foley and Van Valin 1984:333–338; Payne 1997:212).

Finally, some languages have special markings for impersonal subjects, i.e., equivalents of the English neuter personal pronoun *it*, as in *it is raining*.

Object Person and Number

Far fewer languages mark object person and number inflectionally than subject person and number. However, object person and number marking is not uncommon cross-linguistically, and of course is typical (but not definitional) of polysynthetic languages. Generally, when both subject and object persons are marked by inflection, not all logically possible subject and object person and number distinctions are made. This is particularly true for number distinctions.

Aside from these important caveats, all we have said about subject person and number distinctions also applies to object person and number distinctions.

Indirect object person and number that are distinguished inflectionally from direct object person and number are quite rare cross-linguistically. Thus languages that use inflection to indicate subject person and object person distinctions only rarely indicate distinctions between direct object person and indirect object person inflectionally as well.

Associated Motion

Associated motion refers to categories such as andative ‘go and Verb’, venitive ‘come and verb’, cislocative ‘motion toward the speaker’, and dislocative or translocative ‘motion away from the speaker’. The most recent and detailed description of a language with such categories is O’Connor (2007), regarding Lowland Chontal, a language of Mexico.

Grammatical Categories Characteristic of Clauses or Sentences

These include the following:

- Validationality
- Evidentiality

- Mirativity
- Polarity
- Speech act distinctions

Validationality

Validationals are epistemic mood markers. Indeed, validationals are epistemic by their nature: they are to a clause or sentence what an epistemic modal is to a verb or verb phrase. They indicate how sure the speaker is of the proposition. Validational distinctions include dubitative, expressing doubt; potential, expressing possibility; and certain or affirmative, expressing certainty. Validational distinctions can be inflectional, but are probably more often indicated by clitics or separate words. See Payne (1997:251) for a discussion of Validationals.

Evidentiality

Evidentials indicate the source of a proposition, such as visual evidence, hearsay, general knowledge, knowledge from inference, and so on. Evidentials do not, theoretically, indicate how certain the speaker is of a proposition. The first collection of articles on this subject was Chafe and Nichols (1986). That evidentials refer to information source is reflected appropriately in the volume subtitle – "the linguistic coding of epistemology" – since epistemology is the branch of philosophy concerned with how human knowledge is acquired. Other discussions of evidentiality are the short accounts in Chung and Timberlake (1985:242–244), Payne (1997:251), Whaley (1997:224–225), Timberlake (2007:317–318), and the collection in Aikhenvald and Dixon (2003). The most thorough typological survey of evidentiality is Aikhenvald (2004b).

A basic distinction concerns direct versus indirect evidentials. Direct evidentials are used to mark events that have been eyewitnessed, heard, or otherwise directly experienced. Indirect evidentials can be: Inferential (knowledge deduced on the basis of immediate physical evidence), Quotative/reportative (knowledge from hearsay or reported speech), Narrative (knowledge gained from traditional narrative or storytelling), and Gnomic (knowledge considered to be a general truth).

Mirativity

Miratives, occasionally called admiratives, mark knowledge that comes as a surprise to the recipient of it, or knowledge gained unexpectedly (Delancey 1997, 2001).

Validationality, Evidentiality, and Mirativity are easily and often confused. Although they are logically separate, they often need to be analyzed together as one structural system.

Furthermore, these three categories often interact with the TAM system (see Section "Grammatical Categories Characteristic of Verb Phrases"). For example, Miratives and Evidentials often co-occur with past tenses or perfective aspects, since knowledge gained through personal experience or unexpected knowledge is typically about past events; likewise, dubitative Validationals often co-occur with futures, since one can never be too certain of what will happen in the future. Levin (2007) is a recent study of the considerable overlap between tense and evidentiality in Aymara, a Jaqi language of the Andes.

Polarity

Polarity, as applied to the verb phrase, was discussed in Section "Grammatical Categories Characteristic of Verb Phrases". Sentence negation is not uncommonly marked inflectionally on the main verb of the sentence or by an added auxiliary verb. Both word negation and sentence or phrasal negation can be marked syntactically, through the addition of separate words or clitics, and/or through changes in constituent order. The scope of the negative marker needs to be identified.

Under sentence negation, one also needs to pay attention to the common occurrence of special imperative negatives or prohibitives. For example, in Ayacucho Quechua of Peru, the regular negative adverb is *mana*, but the prohibitive adverb is *ama*.

Speech Act Distinctions

Grammatical categories encoding speech act distinctions are often subsumed under Mood/Mode/Modality (discussed in Section "Grammatical Categories Characteristic of Verb Phrases"), but speech act distinctions are really pragmatic rather than syntactic (Levinson 1983:226–283), and they pertain to the whole clause or sentence, rather than just the verb or verb phrase.

The two major types of Speech Act are: Declaratives (also called Indicatives), and Non-declaratives. The two major types of Non-declaratives are Interrogatives or Questions, and Imperatives or Commands (Payne 1997:294, 303, Whaley 1997:233–244).

Types of Interrogatives are the following:

- Yes/no questions (also called nexus questions)
- Tag questions
- Alternative questions
- Question-word questions (also called information or content questions)

Surveys of question types can be found in Ultan (1978a) and Bickford (1998:231). Questions can be inflectionally marked on the verb, but are more often marked syntactically, by separate words or clitics, or by constituent order patterns. It is necessary to identify what the question marker has scope over. In addition to identifying the obvious ways in which interrogation is indicated, the fieldworker would also benefit from looking at the interaction of indefinites and interrogatives, since indefinites and interrogatives are often formally related.

Subtypes of Imperatives are the Jussive, indicating an order, and the Hortative, indicating a suggestion, as in English *let's* .... Some languages have forms for Instructions and Requests. Imperatives can be inflectionally marked on the verb, but are very often marked syntactically, by separate words or clitics, or by constituent order patterns. A typological survey of imperative constructions is in Xrakovskij (2001).

Minor types of Non-declarative Speech Act Distinctions are Exclamations, Interjections, Imprecations, Performatives, Wishes, Vocatives, and expressions of greeting and leave taking. Most of these are not inflectional, the obvious exception being the vocative case of Indo-European languages. Rather, the minor Speech Act Distinctions rely on a variety of routinized syntactic patterns, short words, or interjections.

Politeness strategies typically affect the use of Non-Declarative Speech Act distinctions. For example, indirect speech is often used in place of direct speech to mitigate the force of an utterance (Brown and Levinson 1987). Male and female speech often differs in the marking of Non-Declarative speech acts or indirect speech acts, and in politeness strategies.

Surveys of speech act distinctions are in Sadock and Zwicky (1985), and König and Siemund (2007).

11.3 Morphological Typology and Terminology

11.3.1 Preparatory Reading for Morphological Fieldwork

In this section we focus on morphological fieldwork. The prerequisites for effective fieldwork are once again: typological knowledge, mastery of field elicitation methods, and the ability to develop, reconsider, and revise linguistic analyses.

Regarding these prerequisites, we suggest a study of Nida (1949). Though published more than 60 years ago, this book remains valuable as an extremely detailed introduction to morphological analysis, description, and fieldwork techniques. Six principles are provided for morpheme identification (Nida 1949:7–58). Well worth reading are the statements of principles on which the meaning of morphemes can be determined (Nida 1949:151–157). The book also includes instructions on how to write a description of the morphology of a language (Nida 1949:222–281). It is particularly valuable for its very large set of morphological analysis exercises. There is a bias towards the easiest type of morphology, i.e. agglutinative, and also a bias towards the analysis of indigenous American languages, particularly of Mexico.

Naturally, older works such as Nida (1949) do include some dated concepts and terminology. For example, he does not provide a satisfactory treatment of clitics since he does not distinguish between the phonological word and the syntactic word (Nida 1949:97). The outdated concept of juncture, which never took off anywhere outside the neo-Bloomfieldian school, is also discussed (Nida 1949:102–103). This concept was a result of the fear of “mixing levels”. The idea was that the level of phonology should be analyzed first, the level of morphosyntax should be analyzed next, and morphosyntactic information should not be used to help in the phonological analysis. This idea was first debunked by Pike (1947, 1952). It is now clear to almost everyone that phonology and morphosyntax are not “levels”, but rather modules, and as a result the concept of juncture is no longer needed.

Modern introductions to morphology are Bauer (2003), Aikhenvald (2007), and Haspelmath and Simms (2009). Bauer (2004) is a glossary of morphological terminology. A survey of morphological theories is Spencer and Zwicky (1998), which includes morphological sketches of individual languages. Generative approaches to morphology worth a fieldworker's look are in Aronoff (1976, 1994), Selkirk (1982), and Anderson (1992). Functionalist approaches to morphology are in Bybee (1985) and Bybee et al. (1994).

11.3.2 Formal Morphological Typology

In a grammatical sketch, as in a full-fledged grammar, the fieldworker will need to define the formal morphological typology of the language.

One basic distinction – perhaps more morphosyntactic than morphological in nature – is between analytic versus synthetic languages. In analytic languages, the correspondence of morphemes to words is equal to or close to one-to-one. In synthetic languages, the correspondence of morphemes to words is many-to-one (Sapir 1921:127–128; Payne 1997:27–28; Whaley 1997:128–133). Extreme cases of synthetic languages are called polysynthetic or holophrastic (Fortescue 1994; Evans and Sasse 2002; Mattissen 2004; de Reuse 2006, 2009), and extreme cases of analytic languages, such as Kalam of Papua New Guinea (Pawley 1987, 1993), are called polyanalytic. While there appear to be languages with very little morphology, all languages have syntax; furthermore, the proportion of morphological complexity versus syntactic complexity can vary widely.

Another basic morphological typological parameter distinguishes isolating, fusional, and agglutinative languages. In isolating languages, morphemes are separated by word boundaries; thus the sets of analytic and isolating languages overlap. In fusional languages, information associated with several different inflectional categories is merged together within a single morpheme; in agglutinative languages, every inflectional category corresponds to an independent morpheme or set of morphemes. In other words, it is hard to distinguish morpheme boundaries in fusional languages, but it is easy to segment morphemes in agglutinative languages (Sapir 1921:129–144; Payne 1997: 28–29; Whaley 1997:133–136).

11.3.3 Terminological Issues in Morphology

We can distinguish between complex terminological issues in morphology, for which people cannot even agree on whether the disagreements are formal or functional, and relatively simple issues in terminology, which can be divided into formal morphological terminology and functional morphological terminology.

Complex terminological issues are discussed in Section 11.3.3.1, simpler issues of formal terminology are discussed in Section 11.3.3.2, and simpler issues of functional terminology are discussed in Section 11.3.3.3. Of course, simple and complex are relative terms, and many books have been written about the issues we refer to as "simple". For the fieldworker, the main difference is that the complex terms must be carefully and precisely defined in every description, whereas for the simpler terms it is often, but not always, possible to simply mention them in the description without bothering to define them.

11.3.3.1 Complex Terminological Issues in Morphology

The terms *word*, *morpheme*, *stem*, *root*, *theme* and *base*, are used inconsistently and confusingly in the literature. If the fieldworker uses any of these terms, they need to be defined in every single description.

Word

The fieldworker will discover that, even though most speakers have some concept or idea of what a word is, s/he will have to define the word in the description – both phonologically and syntactically – and will have to carefully point out any mismatches between the phonological word and the syntactic word. Most speakers, and many fieldworkers, will confuse phonological and syntactic words.

There are cases where one can have one phonological word and two syntactic words, which is the phenomenon called cliticization. Cliticization is thus one type of mismatch between phonological word and syntactic word.

Formally, cliticization can look like affixation, in that the cliticized element appears to attach to a word just as an affix does (though in fact, of course, the clitic can be attached to and have semantic scope over a word, a phrase, a clause, or a sentence). A proclitic attaches to the beginning of the domain over which it has scope, and looks formally like a prefix; an enclitic attaches to the end of the domain over which it has scope, and looks formally like a suffix. The clitic equivalent of infixes – i.e., endoclitics – have been described (Harris 2002), but this is much rarer and is still somewhat controversial. The formal equivalent of a circumfix could be called a "periclitic", we suppose.

But there are also cases of cliticization, particularly as fast speech phenomena, where it is not so clear what part of the phonological word is stemlike, and what part of the phonological word is affix-like. Whereas in the case of *do* and *not* becoming *don't* we can treat *=n't* as a clitic to *do*, it is not clear to us what is cliticized to what in the expression *cannot*. Is this *can=* a proclitic to *not*, or is the *=not* an enclitic to *can*? Basic references on cliticization, which do not necessarily agree with our conception of cliticization, include Klavans (1985), Sadock (1991:48–77), Zwicky and Pullum (1983), and Zwicky (1985).

Another sort of mismatch between phonological and syntactic word is the converse of cliticization. These are cases where two phonological words correspond to one syntactic word, discussed in Bickel and Nichols (2007:191–192) and exemplified in the Tibeto-Burman language Lai Chin by Bedell (1997). In some traditions of Sino-Tibetan linguistics, such phonological words that match one syntactic word are called particles (Randy LaPolla, p.c.), but since the term "particle" can mean other things as well (see Section 11.5), we do not recommend this usage.

Since it would take an entire book to properly discuss the issue of what a word is, we recommend reading the latest survey of the issue, Dixon and Aikhenvald (2002), and the discussion on grammatical and phonological words in Dixon (2010b:1–36). For the segmentation of words, Himmelmann (2006b:254–258) is a useful summary.

Morpheme

The fieldworker will find out that most speakers have no concept or idea of what a morpheme is. As in the case of the word, there is considerable disagreement among linguists over the precise nature of the morpheme. Is the morpheme the item in an item-and-arrangement view of morphology, or the item in an item-and-process view of morphology (Hockett 1954)? Or, to put it succinctly, are morphemes entities, or are they rules (Bickford 1998:179)? As first shown by Sapir (1921:57–81), and as formalized in the latest generative literature as the word and paradigm model (Anderson 1992), there is considerable evidence for the morpheme as a rule or process. See Section 11.3.1, for the most recent readings on what a morpheme can be.

For practical purposes, it is best to start with the definition of a morpheme as a linguistic object, something like a minimal unit of meaning equal to or smaller than a word. This works very well with agglutinative languages, but not so well with isolating languages. Furthermore, in a language with extensive fusion, morphemes are best described in terms of a word and paradigm model. The fieldworker should remain agnostic regarding this issue, and adopt the view of the morpheme which best fits his/her analysis of the data. Interesting problems in the description of paradigms are discussed in Himmelmann (2006a).

It is also useful to remember that the American distinction between grammatical morpheme and lexical morpheme corresponds to the French distinction between morpheme and lexeme, and that what Americans call a morpheme is a moneme in much of French linguistics (Martinet 1964).

Stem, Root, Radical, Theme, and Base

The terms stem, root, radical, theme, and base are even more of a terminological can of worms for the descriptive fieldworker. We attempt to show, in Table 11.1, how these terms have been used in three different descriptive grammatical traditions. There are undoubtedly other traditions that use these terms in still other ways.

For Nida (1949:82–83), the root is the nucleus or core of the word, and the stem is what remains of the word after the linguist has taken off all inflection – that is, the stem can include derivational affixes. Both stems and roots can be either bound or

Table 11.1 Comparison of the terms stem, root, radical, theme, and base in three descriptive traditions

	Indo-Europeanist tradition	Eskimologist tradition	Athabascanist tradition	Our remarks
Stem	Word minus inflectional and thematic affixes	Word minus inflectional and derivational suffixes	Word minus inflectional, derivational, and thematic affixes	Best identified negatively, in terms of what is removed
Root	Semantic core of a word	Semantic core of a word	Abstract representation of a stem set	“Root” is best viewed as a reconstructed historical concept
Radical	Same as root	Not used	Not used	Best viewed as a synonym for root
Theme	Word minus inflection	Not used	Word minus inflectional and derivational affixes	A term used only if a set of thematic affixes is recognized
Base	Not used	Word minus inflection	Word minus inflection	

free. All stems contain at least one root, but not all stems are roots. This is basically the Indo-Europeanist conception of stems and roots, with the exception that Nida ignores the difference between stems and themes. In the Indo-Europeanist tradition, themes refer to stems containing a thematic element; a thematic element is an element that has to be present although its function and meaning are not always clear to either the speaker or the linguist. The short definitions of root and stem in Payne (1997:24) basically agree with Nida’s. Interestingly, the more typologically oriented texts by Whaley (1997:290) and Croft (2003:255) discuss roots, but not stems.

11.3.3.2 Formal Morphological Terminology

Here we discuss formal terms which are fairly straightforward and uncontroversial. Morphology can be of two types: Concatenative and Non-concatenative. Concatenative morphology orders morphemes like beads on a string, or like links of a chain, i.e., one after another. Of course, in concatenative morphology, the morphemes can also be structured hierarchically with respect to one another, the way words are organized in syntax (Section 11.4.4). Any morphological process that does something different from concatenation is non-concatenative.[8] While every language has concatenative morphology (if it has morphology at all), not all languages have non-concatenative morphology.

[8] Concatenative is not the same as what is called additive by Nida (1949:69), because he considers reduplicatives (see below) to be additives, whereas reduplicatives are not concatenative. Concatenative is also not the same as what is called linear in Bickford (1998:177), because he considers infixation to be nonlinear, whereas infixation is concatenative.

Concatenative Morphology

Concatenative morphology includes affixation and compounding. Affixation is the concatenation of a bound element to a stem; we can distinguish:

- Prefixation, i.e. the concatenation of an affix preceding the stem
- Suffixation, i.e. the concatenation of an affix following the stem
- Infixation (or discontinuous affixation (Bickford 1998:177)), i.e. the insertion of an affix inside the stem
- Interfixation or Interposition, i.e. the insertion of an affix in between two stems, so that the affix never occurs without a stem following and preceding it (Helmbrecht and Lehmann 2008:279–280; Bickel and Nichols 2007:199–200)
- Simulfixation, i.e. the concatenation of an affix at two positions on the stem; the affix has only one meaning, and its two parts never occur separately. The most common subtype of simulfixation is Circumfixation, the concatenation of an affix preceding and following the stem, or in other words, a prefix and a suffix always occurring simultaneously (Bickel and Nichols 2007:200–201)

One issue related to concatenative morphology has to with affix ordering. This is discussed in Grimes (1967), Muysken (1986), and Bickel and Nichols (2007:214–220).

In addition to surface ordering, concatenative morphology often exhibits a layered or hierarchical structure. When there does not appear to be a layered or hierarchical structure, we talk about Templatic, Slot-and-filler, or Position class morphology. A template is not the preferred mode of describing a sequence of affixes (Rice 2006:252–253), but it sometimes is the best the fieldworker can do (Simpson and Withgott 1986).

Compounding is the concatenation of two stems (or roots) into one word (Anderson 1985a:40–43; Aikhenvald 2007:24–35). Both stems can be free, as in *blackbird*; one stem can be free, as in *cranberry*; or both stems can be bound, as in *telescope*. Compounding will be further discussed under Section 11.3.3.3.

Nonconcatenative Morphology

Subtypes of nonconcatenative morphology are: Suprasegmental, Subtractive, Replacive, and Reduplicative.

Suprasegmental morphology includes suprasegmentals such as contrastive tone, stress, or nasalization over the domain of a syllable or a word.[9]

Subtractive morphology is relatively rare (Nida 1949:75); this morphological process involves deletion of part of the stem. Examples can be found in O'odham (Pima and Papago), a Uto-Aztecan language of Arizona; Polynesian languages; and French.

---

[9] The term "suprafix" used by Nida (1949:69) to describe this phenomenon has not caught on.

Much more common is Replacive morphology, in which a morphological process replaces one or more segments in the stem (Nida 1949:71–72). Replacive morphology is also called Root and pattern morphology, or – in descriptions of Semitic and other Afroasiatic languages – Templatic morphology, or Morpheme internal change (McCarthy 1981). When replacive morphology involves only vowel changes, as is typically the case in Indo-European languages, it is known by the German term Ablaut, or by the terms Mutation or Apophony (Bickford 1998:177). When replacive morphology involves only consonants, it is more often called Consonant gradation rather than Consonant mutation.

In Reduplicative morphology, generally simply called Reduplication, the stem or some part of the stem is copied and then added to that stem as an affix. In Complete Reduplication the whole stem is copied. In Partial Reduplication a phonologically describable piece of the stem (a syllable or part of one or more syllables) is copied. Reduplication is quite common in the languages of South Asia, as shown in Abbi (1992). For a detailed study of Reduplication in a Tibeto-Burman language of India, see Chelliah (1997:274–281). It is less favored in Indo-European languages. Triplication, in which the stem or stem piece is repeated two times and then added to the stem, is much less common, and could theoretically be analyzed as recursive reduplication or as a phenomenon entirely distinct from reduplication (Uray 1954).

Two terms having to do with formal types of morphemes, but which do not really refer to nonconcatenative morphological operations, can profitably be mentioned here: Total fusion or Portmanteau morphology, and Suppletive morphology or Suppletion.

Total fusion or Portmanteau morphology occurs when two morphemes are expected, according to paradigm, but only one is found, in essence covering the meaning of the two expected morphemes (Bickford 1998:177). An example is the Lakota *čhi*- ‘I subject-you (sg.) object’, where one expects a sequence of *ni*- ‘you (sg.) subject’ and *wa*- ‘I subject’.

Suppletion occurs when a stem with a certain amount of morphology is expected, but instead a formally unrelated stem is found (Bickford 1998:162–163). The textbook example is English *was* and *were* for expected *be+d* (past). Other examples are *went* for *go+d* (past), and *is* for *be+s*.

11.3.3.3 Functional Morphological Terminology

We now turn to a discussion of the functional terminology used to describe morphology.

Inflectional and Derivational Morphology

There has been much theoretical discussion regarding the usefulness of the traditional distinction between inflectional and derivational morphology. For purposes of morphological analysis, this distinction is indeed useful. We recommend the classic discussion in Nida (1949:99) and the summaries in Payne (1997:25–26) and Whaley (1997:120–123).

The widely accepted idea that derivational affixes are inner, i.e. closer to the stem, and that inflectional prefixes are outer, i.e. closer to the edge of the word, is certainly attested cross-linguistically, but there are numerous exceptions. In de Reuse's experience, this idea seems most applicable to exclusively suffixing language families, such as Quechuan or Eskimo-Aleut, but the prediction does not work for heavily prefixing language families such as Siouan and Athabascan.

Productivity is the ability of a bound morpheme to occur with any stem to produce a semantically plausible combination. Productivity has been used as a criterion to distinguish between inflectional and derivational morphology, but since levels of productivity are sometimes hard to establish, this is not always a useful approach. A strict, nonfuzzy, non-continuum analysis of productivity – i.e., where morphemes are either productive or non-productive – is preferred, but is not always possible. There are certainly degrees of non-productivity; for example, some non-productive elements occur with thousands of elements, some with only two or three. See Anderson (1985a:16–22), Comrie and Thompson (1985:357–358, 2007:342–343) and especially Aikhenvald (2007:49–58) for useful discussions.

It is best to analyze predictability and productivity together. If the fieldworker can predict that a certain form exists (assuming it is semantically and pragmatically plausible), and it does indeed exist, then the fieldworker should begin to investigate whether the morphological construction is productive. In doing so, it is important to distinguish the productivity of a grammatical category (e.g., singular vs plural) from the productivity of a particular allomorph: English pairs such as *foot/feet*; *tooth/teeth*; *child/children*, *ox/oxen* are definitely not predictable morphology. So within a productive grammatical category, there can be irregularity. In the next section, we first discuss unambiguous cases of derivational morphology, then unambiguous cases of inflectional morphology, and finally cases that seem to be neither derivational nor inflectional.

Derivational Morphology

By all accounts, derivational morphology is morphology that effects a change in lexical category. Derivational morphological terminology can describe the original lexical category; e.g. denominal or deverbal, meaning that the original lexical category must have been a noun or a verb, respectively. Derivational morphological terminology can also describe the end result lexical category after derivational morphology has been added, e.g., verbalization or nominalization. A chart of the terminology in common use is given in Table 11.2. The pattern is clear.

Table 11.2 Terminology of lexical category changing derivational morphology

	End result lexical category: Noun	End result lexical category: Verb	End result lexical category: Adjective	End result lexical category: Adverb
Original lexical category: Noun	–	Denominal verbalization	Denominal adjectivalization	Denominal adverbialization
Original lexical category: Verb	Deverbal nominalization	–	Deverbal adjectivalization	Deverbal adverbialization
Original lexical category: Adjective	Deadjectival nominalization	Deadjectival verbalization	–	Deadjectival adverbialization
Original lexical category: Adverb	Deadverbial nominalization	Deadverbial verbalization	Deadverbial adjectivalization	–

A survey of nominalization is in Koptjevskaja-Tamm (1993). Comrie and Thompson (1985, 2007) make the following distinctions in deverbal nominalization:

- Action or Activity nominalization, resulting in Action or Activity nouns
- State nominalization, resulting in State or Stative nouns. See also Noonan (1985:108) and Payne (1997:224–230) on this point
- Participant nominalization, which can be further subclassified into:
  - Agent nominalization, resulting in Agent or Agentive nouns
  - Patient nominalization, resulting in Patient or Patientive nouns
- Instrument nominalization, resulting in Instrumental nouns
- Manner nominalization, resulting in Manner nouns
- Location nominalization, resulting in Locative nouns
- Objective nominalization, resulting in nouns designating the result or product or cognate object of an action
- Reason nominalization, resulting in Reason nouns

Denominal verbalization is illustrated in Comrie (1985b:346–348). Useful subtypes are denominal verbalization from predicate nouns and denominal verbalization from arguments of predicates, often called light verbs (in a different sense from the auxiliaries discussed in Section 11.2.3.1). A useful online reference to the literature on light verbs is http://mwe.stanford.edu/lightverbs.html. Denominal light verbs typically have one of the following meanings: 'to be N'; 'to have N'; 'to produce or manufacture N'; 'to use N (instrument) to V'; 'to go to N (destination)'; 'to spend the N (period of time such as winter, night, etc.)'; and 'to feel N (emotion)' (Comrie 1985b:346–348).

Deadjectival verbs are discussed and illustrated in Comrie (1985b:345–346). It seems to be cross-linguistically true that adjectivalizations and especially adverbializations are less common than nominalizations or verbalizations, and that deadjectival and especially deadverbial derivation is much less common than denominal and deverbal derivation.

There are also derivational processes where something other than the lexical category is changed.

An example is verb-to-verb derivation -- a process, which in effect, would fill in the gap under "deverbal verbalization" in Table 11.2! As discussed and illustrated in Comrie (1985b), verb-to-verb derivation is often a valence (or valency) changing process (see Section "Grammatical Categories Characteristic of Verb Phrases"). Valence can be increased, i.e., adding obligatory arguments to the verb, or decreased, i.e., removing obligatory arguments from the verb. Valence increasing derivational processes are applicatives (Comrie 1985b:316–318) and causatives (Comrie 1985b:323–325). Valence decreasing derivational processes include anti-causatives. These are also called unaccusatives, particularly when zero morphology is involved (Perlmutter 1978). Other valence decreasing derivational processes include reciprocals and reflexives (Comrie 1985b:325–327).

The fieldworker should keep in mind that valence changing processes are not always a result of derivational morphology, as shown in Section "Grammatical Categories Characteristic of Verb Phrases", and discussed under "Inflectional Morphology".

Noun-to-noun derivation (i.e. "denominal nominalization") is discussed and illustrated in Comrie and Thompson (1985:395–396, 2007:379–381). Noun-to-noun derivation typically involves the derivation of abstract nouns from concrete nouns, as in *childhood* from *child*, or the derivation of diminutive, augmentative, or pejorative nouns, as in *duckling* from *duck*.

Verb-to-verb derivation and noun-to-noun derivation are most developed in polysynthetic languages like Eskimo (de Reuse 1994).

Further interesting examples of derivational morphology are given in Anderson (1985a:24–39) for Kwakw'ala (Kwakiutl of the American Northwest) and for Classical Arabic.

Inflectional Morphology

We now discuss inflection, and inflectional categories. Most schools of linguistics will agree, following Anderson (1982), that inflectional morphology is morphology that cannot be understood without reference to syntax. The clearest examples are case marking and agreement morphology. However, since there are languages where case marking and/or agreement is conveyed by syntactically independent words, case marking and agreement are not necessarily morphological issues. A survey of agreement is in Corbett (2006).

Useful surveys of inflectional categories are found in Anderson (1985b), Bickford 1998:133–134), and Bickel and Nichols (2007). Haspelmath (1996) argues that one must accept the existence of word-class-changing inflection. But since we consider word-class changing to be a definitional feature of derivational morphology, we are not convinced that positing this same characteristic for inflection is a useful approach; instead, it is possible that such cases should be considered to be neither inflection nor derivation.

Anderson (1985b:172–174) posits an interesting classification of the inflectional categories associated with Nouns, Verbs, Adjectives, and Adverbs into Inherent, Relational, and Agreement Inflection. Inherent inflectional categories depend only on the semantics or pragmatics of the inflected word; relational inflectional categories can only be understood in terms of syntactic relationships with other words; and agreement inflectional categories can arise through copying of an inherent or relational category from one word onto another.

For nouns, inherent categories include: number, gender (which can be based on sex or animacy) or noun class, size (diminutive or augmentative), referentiality/definiteness/specificity, and deixis (Anderson 1985b:174–179). Relational categories include case and possession (Anderson 1985b:188–189).

For verbs, inherent categories include: tense/aspect/mood, evidentiality/mirativity/validationality, and polarity (Anderson 1985b:189–191). Parameters such as "conjugation class" can be considered inherent categories, but it is not always clear that these are inflectional in the first place. Relational categories for verbs can include: person and number of subject, person and number of object, switch reference, subordinated/subordinating status, and – assuming that they are not derivational in the language in question – valence changing categories such as voice, causativity, reflexivity, reciprocity, middle, benefactive, and possessive (Anderson 1985b:191–194),

For adjectives, the inherent category typically involves comparison: i.e., equative ("as big as"); comparative ("bigger than"), and superlative ("the biggest"). A survey of comparison is in Stassen (1985). Relational categories, relatively rare for adjectives, are exemplified by the inflection marking modifier versus predicative status of the adjective, as seen in Dutch. In the Dutch phrase *de mooie vrouw* 'the beautiful woman', the adjective *mooi* 'beautiful' is inflected with the suffix *–e* to mark its modifier status. However, in *de vrouw is mooi* 'the woman is beautiful', the suffix *–e* is absent, marking its predicative status.

Theoretically, any of the above could also be Agreement categories, but the most common types of agreement involve person, number, gender or noun class, referentiality/definiteness/specificity, and case (Anderson 1985b:194–200).

For adverbs, the inherent and relational categories, if they occur at all, are typically like those of adjectives. Anderson (1985b:200–201) points out that the kinds of agreement associated with adverbs are so similar to those associated with verbs that we might well doubt the existence of a category of adverbs distinct from verbs in such languages.

Some Dutch dialects exhibit agreement between the modifier status of an intensifying adverb and a following modifying adjective. In Standard Dutch, the sequence [Adv[Adj[Noun]]] is inflected as [Adv[Adj_{MOD}[Noun]]], where the subscript MOD marks the modifier status inflectional ending of the adjective; but in these other Dutch dialects, this MOD agrees with a preceding adverb as well, and the resulting construction is [Adv_{MOD}[Adj_{MOD}[Noun]]]. So, to reuse the example from a previous paragraph, in these Dutch dialects, *de heel mooie vrouw* 'the very beautiful woman' would be *de hele mooie vrouw*.[10]

[10] Orthographically, the double *ee* is reduced to one *e* in *hele*, but phonologically, we have *heel* plus the inflectional ending *–e*, indicating modifier status.

There are also interesting paradigmatic dependencies between grammatical systems. These dependencies are particularly important in understanding inflectional systems. These are cases where the choices available in inflectional system X depend on the choice that is made in inflectional system Y. To take a simple example, the choice between singular and plural in English subject pronouns depends on the choice made for person. Choices between singular and plural are possible for all pronouns, except for the second person, where no choice is possible, since the form *you* covers singular and plural. The foundational article on such dependencies is Aikhenvald and Dixon (1998), and a summary is in Dixon (2010a:162–165).

Morphology That Is Neither Derivational nor Inflectional

Some morphology may not be either clearly inflectional or clearly derivational. These instances are sometimes analyzed as functional types on a cline or continuum between inflectional and derivational morphology (Bybee 1985).

Another approach, developed specifically to address polysynthetic morphology that cannot be clearly classified as inflectional or derivational, is represented by de Reuse's (2006) postulation of "productive noninflectional concatenation" (PNC). Mel'chuk and Beck (2006) distinguish a quasi-inflectional morphology, which is a similar idea. A chart summarizing the features distinguishing inflection, derivation, PNC, and syntax, taken from de Reuse (2009) is given in Table 11.3. Note that features which typically identify derivational morphology, here distinguish (non-productive) derivation from PNC, and that PNC has much in common with syntax.

Headedness in Morphology

The difference between heads and dependents was defined in Section 11.2.2. A different matter is whether a morphological construction is headed, and if so, where the head is located. There is a debate within morphological theory about what is the head in stem+derivation or stem+inflection constructions. In some cases, it can be argued that the stem is the head, in others, that the derivational or inflectional element is the head, and in others that the construction has no head (Zwicky 1985).

Table 11.3 Criteria of inflection, derivation, PNC, and syntax

	Inflection	(Nonproductive) derivation	PNC	Syntax
Productive?	Yes	No	Yes	Yes
Recursivity possible?	No	No	Yes	Yes
Necessarily concatenative?	No	No	Yes	Yes
Variable order of elements possible in some instances?	No	No	Yes	Yes
Interaction with syntax possible?	Yes	No	Yes	Yes
Lexical category changing possible?	No	Yes	Yes	Yes

In compounding, it is easy to determine headedness. Compounds are endocentric if they are headed, and exocentric if they are headless. An example of an exocentric compound is the English word *bigshot*. Further useful discussion of compounding is in Hoeksema (1985).

The literature on Mandarin Chinese compounding provides a useful functional typology of compounding. Compounds can be classified as endocentric, (Modifier-Modified in Chinese), or exocentric which can be subclassified as Verb–object, Subject–predicate, Coordinate, and Resultative verb compounds (Anderson 1985a:43–52; Aikhenvald 2007:33–34). A more extensive typology of Mandarin Chinese compounds is in Chao (1968:359–480), and in Li and Thompson (1981:45–83). Marchand (1969) contains an extensive functional classification of compound types for English.

There are various other semantic types of headless compounds, particularly for East Asian and Southeast Asian languages, such as compounds of two synonyms, or compounds of two antonyms, something like "hot-cold" meaning 'temperature'.

Noun incorporation is often, but not always, a type of endocentric compounding of a noun stem and a verb stem, the verb stem being the head (Mithun 1984; Gerdts 1998; de Reuse 1993).

Verb serialization is also sometimes a type of endocentric compounding of two (or more) verb stems, one of the verb stems being the head (Aikhenvald and Dixon 2006). More on verb serialization as a syntactic phenomenon is in Sections 11.4.3 and 11.4.5.

Grammaticalization and Lexicalization

In this discussion of functional issues in morphology, a few comments on the terms lexicalization and grammaticalization are in order. In the linguistics literature, lexicalization refers to the historical evolution of a grammatical element (either a morpheme or a grammatical word) into a lexical item, i.e., a free morpheme with lexical meaning. Conversely, grammaticalization refers to the historical evolution of a lexical item, into a grammatical element (either a morpheme or a grammatical word) (Hopper and Traugott 1993).

Some formalist linguists, such as Newmeyer (2001), object to the concept of grammaticalization as a distinct historical process. We agree that the concept of grammaticalization can be abused, and to some extent the debate is a terminological one, which means that the fieldworker should be cautious with this concept. In some cases, it is probably a good idea to talk about morphologization or syntactization as subtypes of grammaticalization, or even of derivationalization or inflectionalization as subtypes of morphologization.

In the functionalist and cognitive linguistic literature, the terms grammaticalization and lexicalization are also used to refer to concepts of synchronic grammar. That is, when a particular semantic distinction is marked in the grammar it is called grammaticalized, and the fact that it is present is called grammaticalization. Conversely, when a semantic distinction is marked in the lexicon but not marked in the grammar, it is called lexicalized, and the fact that it is present is called lexicalization (Bybee 1985; Bybee et al. 1994). Both usages have respectable pedigrees, but the fieldworker should not confuse diachronic grammaticalization/lexicalization with synchronic grammaticalization/lexicalization.

Causatives are a "locus classicus" of grammaticalization versus lexicalization in the cognitive or functional sense, as summarized in Comrie's (1985b:330) discussion of analytical causatives (i.e., through syntax) ("cause to die"), morphological causatives (i.e., through derivational morphology, occurring in many languages), and lexical causatives ("kill").

Other fruitful work related to the issue of grammaticalization versus lexicalization in the cognitive or functional sense is Vendler's (1967) classification of predicates, and elaborations thereof (Chung and Timberlake 1985:218; Timberlake 2007: 284–288; Wilhelm 2007:1–22). This classification can be used to understand the differing morphological and syntactic characteristics of identified predicates. For instance, the classification helps with the description of Aktionsart (or inherent aspectual value, or lexical aspect) (see Section "Grammatical Categories Characteristic of Verb Phrases"). While it is clear that all languages can distinguish telicity and durativity as semantic features, the fieldworker should remain cautious in distinguishing Aktionsart from derivational or inflectional aspect.

The most recent work on grammaticalization and lexicalization in the cognitive or functional sense is that of Talmy (1985:126–127, 2007), who provides a table of 35 semantic categories and their lexicalization patterns, explained within a cognitive framework.[11] This table should be of direct use to the fieldworker.

11.4 Syntactic Typology and Terminology

11.4.1 Preparatory Reading for Syntactic Fieldwork

Necessary preparation for syntactic fieldwork is the reading and study of the vast literature on syntactic analysis and typology, both rich and complex fields. We cannot in this section review all existing accounts of syntactic typology, but the literature in Section 11.2.1 covers syntactic typology fairly well. For syntactic analysis, a useful guide is Green and Morgan (2001). An introduction to syntactic concepts and terminology is in Luraghi and Parodi (2008).

What we attempt in Section 11.4 is a broad-stroke vademecum of the syntactic terms and concepts most useful to fieldworkers, trying to distill the best in recent

[11] The fieldworker would do well to study this approach, while also being aware that, for purposes of morphological analysis, Talmy's distinctions of "root", "satellite", and "inflection" as primitives are not precise enough, since "satellite" is a cover term for a particle or a productive derivational element. A more expanded treatment of Talmy's cognitive semantics is Talmy (2000).

syntactic theories, with suggested references. The sections are organized as follows: Theories of Syntax (Section 11.4.2); Ordering, Transforming and Combining (Section 11.4.3); Constituents, Constructions, and Hierarchies (Section 11.4.4); Headedness and Dependencies (Section 11.4.5); Modularity (Section 11.4.6); Syntactic Mechanisms (Section 11.4.7); and Fuzziness, Clines, and Grammatical Hierarchies (Section 11.4.8).

Also, the fieldworker will find that there is no clear boundary between syntax and discourse. All syntacticians and discourse analysts will agree that there is such a thing as a clause, and that there are levels of structure above the clause. Are such levels discourse levels or syntactic levels? Is the sentence a discourse level, a syntactic level, or both? Or are syntax and discourse not even levels in the first place? In his/her analysis, the fieldworker will have to be mindful of the fact that there are excellent arguments for all of these theoretical positions. Useful surveys of discourse analysis are in Schiffrin (1994), Dooley and Levinsohn (2001), and Kärkäinen et al. (2007). The latest studies of the relationship between discourse and grammar are in Mushin and Baker (2008). Studies of the relationship between discourse and pragmatics are in Goddard (2006).

11.4.2 Theories of Syntax

We agree with Gil (2001:125–128) that there is no such thing as description without theory, and there is no such thing as theory without description. However, as fieldworkers, we can also be sympathetic with Abbi's (2001:3) view that "theory-dependency more often binds the fieldworker's hands and restricts her/his vision."

The issue depends, of course, on what we mean by "theory" and by "theory -dependency".

For the descriptive linguist, as far as "theory" is concerned, we find Dryer's (2006) position the most useful. Dryer posits a distinction between descriptive theories and explanatory theories, which can be traced back to Chomsky (1965:24–27). Dryer points out that Chomsky quickly moved towards the idea that a single theory can be used as both a descriptive theory and an explanatory theory, and he argues that this need not be so (2006:207–212). In other words, the descriptive fieldworker should work within a descriptive theory and not worry about whether the theory is also explanatory. Dryer also argues that BLT (Basic Linguistic Theory; Section 11.2.1.1) is a good descriptive theory (2006:210–212, 225–227). We will return to BLT below.

Furthermore, the distinction between descriptive theory and explanatory theory reminds one, of course, of Chomsky's earlier evaluations of a grammar in terms of "observational adequacy", "descriptive adequacy", and "explanatory adequacy" (Chomsky 1964:63). What happened to observational adequacy? Explanatory theories, regardless of whether they are functional or formal/generative, have tended to pay far less attention to observational adequacy, than did earlier theories such as American structuralism. This lack of regard for – or even contempt for – observational adequacy on the part of some explanatory theoreticians has been the cause of satiric critical essays by Postal (1988) and Sadock (1996), and not undeservedly so. It is therefore important that the descriptive linguist work within a theoretical framework that will not force him/her to compromise on the standards of observational adequacy.

Let us now turn to the often misunderstood issue of "theory dependency". The fieldworker needs to be aware that much of the literature on a particular topic is grounded in particular theoretical frameworks and that descriptions of linguistic phenomena written in a particular framework are shaped by its theory-internal deductions and presuppositions. The term "theory-internal" is not intended here as a pejorative one, since by definition, all elements of a theory are theory-internal.

It is just that within any theory some deductions are made which are (1) more abstract, (2) less elegant, and (3) harder to translate into other theoretical frameworks than others (Evans and Dench 2006:7). For lack of a better term, we will call these "theory-internal" deductions. Such deductions are the elements that the fieldworker should be cautious about. It is not that they are wrong, but rather that they might not be the most efficient in the fieldworker's pursuit of descriptive adequacy.

Moreover, theories of linguistics, and in particular theories of syntax, are particularly good at elegantly accounting for some phenomena, while ignoring or discounting others. Every theory of syntax does this to various extents and in different ways. This means that, in his/her theoretical orientation, the fieldworker has to be eclectic in order to make best use of different theories. So when we say that the descriptive fieldworker should not be too be "theory-dependent", what we mean is that s/he should be cautious of certain theory-internal deductions or presuppositions, and that s/he should be eclectic to some extent.

Furthermore, some useful syntactic concepts do not appear to be theory-internal in any disturbing way, but nevertheless give every theory some trouble. Let us take, as an example, the concept of grammatical relations.

Grammatical relations is the term for basic syntactic functions as they relate to a verb, such as "subject (of)", "object (of)", "indirect object (of)". Whereas in some languages there is very little evidence for such grammatical relations, grammatical relations have been so useful that it is best not to consider them theory-internal. However, grammatical relations are the textbook case of a concept that every theory of grammar considers in a different light, but which no theory of grammar can do without. Needless to say, the fieldworker will find them both useful and elusive. It is quite disturbing to a beginning fieldworker that some languages show so little interest in a grammatical relation such as subject.

The literature on subjects is immense; a few useful references are Li (1976) and Mulder (1989). If the specific tests for identifying subjects listed in Andrews (1985:104–109, 2007a:166–179) and Kroeger (2004:212–298) are used, some will fail, but probably not all. So the best the fieldworker can do is point out in his/her description which tests work, and which do not, and tentatively decide what should be called "subject" or "object" on that basis. Maybe grammatical relations are fuzzy categories (Section 11.4.8) in all languages. A good overview of grammatical relations in several theories is Farrell (2005). A survey of grammatical relations is in Palmer (1994), and shorter accounts are in Dixon (1989), Payne (1997:129–132), Whaley (1997:67–73), and Kroeger (2004:14–16).

For all the reasons above, it is important for the fieldworker to be aware of a range of syntactic theories, and of what they can and cannot do (Bowern 2008:10–12). Obviously, this knowledge is necessary for a critical reading of the literature and for assessing which representations are the most useful in describing the target language. Since it is the fieldworker's charge to capture the brilliance of the language under study, it is crucial to find a framework or frameworks that will project that brilliance. Suppose, for example, that a particular theory is simply uninterested in the interaction of pragmatics and case marking, but the language under study does show such an interaction. Rather than ignoring these interesting facts, the fieldworker should be aware of other possible ways of representing them.

It is not entirely clear whether the "non-theory internal" elements of syntactic theory can be brought together in what Dixon (2010a) has called "basic linguistic theory" (BLT). As pointed out in Section 11.2.1.1, we are favorably disposed towards the idea of a "basic linguistic theory" but remain agnostic as to whether such a theory can be a coherent whole. We do agree with Dryer (2006) in his argument in favor of BLT, in that we see the goal of the fieldworker to be the production of an elegant and relatively complete syntactic description that forms a coherent whole for the language under study. The grammar should not, however, try to describe "each language entirely on its own terms" (Evans and Dench 2006:5), not only because languages have to be typologically comparable, but also because the result would be difficult to read. The key word in the quotation above is "entirely".

In our own conception of BLT, which is not exactly that of Dixon (2010a), we suggest a period of introspection when the fieldworker assesses which theories fall within acceptable parameters of representing syntactic competence for him/her. Then within these parameters, the fieldworker should pick and choose, in an unabashedly eclectic way, what works best from these theories for the language under consideration.

What works best will be language dependent to some extent, but might include elements from our favorite grammatical theories, presented here, in approximate order of appearance: McCawley's view of syntax (McCawley 1995, 1998), Lexical-Functional Grammar (Kroeger 2004), Role and Reference grammar (Van Valin and Lapolla 1997), Autolexical or Automodular Grammar (Sadock 1991), Construction Grammar (Goldberg 1995; Croft 2001; Fillmore et al. 2003), and Basic Linguistic Theory (Dixon 2010a, b).[12]

Useful model studies of syntactic topics by fieldworkers interested in theory are in Nichols and Woodbury (1985), Foley (1993), and Rice (2006).

[12] We are not convinced that the interesting theory based on universal semantic primitives proposed by Wierzbicka (1986a, 1988, 1994) is easily applicable to descriptive fieldwork.

11.4.3 Ordering, Transforming, and Combining

The fieldworker will observe that words occur in a certain order, that words, phrases and clauses are combined, and that various constructions seem to be somehow related to each other.

One of the first facts about a language that a fieldworker will want to establish early on is what is traditionally called “word order”. What is really meant by this is first and foremost argument order, i.e. the relative order of the subject argument (S), and the object argument (O), with respect to the verb (V) (Greenberg 1963). For example, it can be claimed that OSV is the basic order in a language, as described for Warao (an indigenous language of Venezuela, Guyana and Suriname) by Romero-Figueroa (1985).[13]

A more sophisticated approach to “word order” is that of constituent order typology, in particular the distinction between verb initial and verb final languages. On the basis of this distinction, other patterns of constituent order seem to be predictable with better than chance probability. Basic surveys are Hawkins (1983), Dryer (1997), Song (2001:49–137), and Dryer (2007a), and they are summarized in Comrie (1981, 1988), Hawkins (1989), Payne (1997:71–74), Whaley (1997: 79–95), Bickford (1998:109), Abbi (2001:177), and Croft (2003:52–80). On a comparison between the generativist and typological approaches to constituent order, see Hawkins (1985). For constituent order and discourse, see Downing and Noonan (1995), and for studies on iconicity in constituent order, see Haiman (1985).

The basic questions the fieldworker will ask are the following: Is there such a thing as a basic word or constituent order in this language (Brody 1984; Mithun 1992)? Or, in another tradition, is this a configurational language, sometimes called an X-bar language (Jackendoff 1977), or a nonconfigurational language, sometimes called a W* language (Hale 1983)? For more on how to determine the basic constituent order of a clause and how to determine if constituent order is flexible or strict, see Payne (1997:77–82), Whaley (1997:96–107) and Kroeger (2004:141).

The fieldworker will sometimes come to the conclusion that the language does not have a basic constituent order, but also that it is neither configurational nor nonconfigurational, and instead occupies a pragmatic limbo in between. Illuminating discussions on the problems of describing pragmatically conditioned constituent order are in Gil (2001) for Malay, and in LaPolla and Poa (2006:273–281) for Chinese. Other studies of the relationship between pragmatics and constituent order are Payne (1990, 1992).

Even though the idea of transformational relationships is *passé* for formal (generative) syntacticians, we nonetheless find it a useful descriptive device,

---

[13] We are somewhat ambivalent about the term argument, because that is really a term from predicate calculus (Payne 1997:174), whereas non-argument (and non-obligatory) elements in the clause are called adjuncts (Payne 1997:317; Kroeger 2004:9–12), and we are not sure adjuncts is also a term of predicate calculus. The usefulness of terms such as arguments and adjuncts is that it allows us to talk about the order of meaningful elements in a sentence, without making a commitment to issues of hierarchical structure, which will be discussed in Section 11.4.4.

following the work of Jim McCawley and Háj Ross (p.c.). These devices for relating underlying or core syntactic strings with surface syntactic strings were first posited by Harris (1952, 1957) and incorporated in the early versions of Chomsky's grammatical models (Chomsky 1957, 1965). The fieldworker can remain agnostic about the explanatory implications of transformations and yet still find it useful to represent some constructions using transformational machinery. We find it hard to imagine studying the passive in any theory without in some way relating it to an active voice. Siewierska (1984) and Shibatani (1988) are useful studies of passives, and Kroeger (2004:53–87, 103–134) is a good survey of other structures that can be described in terms of transformations. Other helpful discussions of relationships usefully described in terms of transformations are in Bickford (1998:329), McCawley (1995, 1998), Foley and Van Valin (1984:149–168), and Payne (1997:204–209).

Transformations and word or constituent order are also relevant to clause combining. Surveys of clause combining or clause linkage are in Payne (1997:306) and Bickel (1991). A cross linguistic typology of the semantics of clause linking is in Dixon and Aikhenvald (2009). The terminology regarding clause combining or linkage is complex, is often language-specific, and can be very confusing. In this section, we review the main terms and their definitions.

Clause combining can be carried out through parataxis, coordination, or subordination. Subordination will be discussed in Section 11.4.5.

Parataxis is the juxtaposition of two clauses or phrases without any connecting element. This is much more common in spoken language than imagined, because intonational features not represented in written language make the connection between clauses clear. Parataxis of subject sharing verbs within a verb phrase is called verb serialization. See Noonan (1985:76–82, 2007) and Aikhenvald and Dixon (2006) on serial verb constructions.

Logical relations in coordination can be: conjunction "and"; disjunction "or", or exclusion "a and not b" (Payne 1997:341). Payne (1985b) includes a detailed discussion of coordination in which he claims that it is rare to find disjunction and conjunction indicated by the same syntactic marking; however, homonymy of conjunction and disjunction marking does exist in Siberian Yupik Eskimo. Meaning is elucidated by context. What never occurs, with good reason, is homonymy of conjunction and exclusion.

It is often the case that mechanisms for clause combining are similar to mechanisms for phrase combining. However, the fieldworker should be on the look-out for coordinating conjunctions (e.g. *and*) that can only conjoin phrases and coordinating conjunctions with the same meaning that can only conjoin clauses. A thorough survey of coordination is in Haspelmath (2007).

11.4.4 Constituents, Constructions, and Hierarchical Structure

The fieldworker will, of course, become aware of the organization of constituents into hierarchical structures. In Chomskyan formalisms such as phrase-structure grammar or X-bar theory, this structure is represented by attaching elements to trees. Construction grammar places less of an emphasis on hierarchical structure, but the idea is the same, and elements are represented in boxes or semantic maps.

Hierarchical structure is less present in the consciousness of the native speaker (although easily recognized when pointed out) than surface ordering facts; it is also harder to represent in the fieldworkers' notes. The question of how many phrase structure nodes or constructions one needs is never easy to answer.

The well known phenomenon of recursion in syntax, accepted by all syntacticians as universal in human language with the exception of Everett (2005:630) needs to be built into one's hierarchical structure. Everett's position that recursion need not occur in all human languages has given rise to extensive debate (Nevins et al. 2009a, b; Everett 2009a, b). From the fieldworker's point of view, such a debate does not mean that s/he has to take sides, but it does mean that s/he will have to describe recursion more carefully than has been done in the past, and if recursion does not seem to occur, s/he must document that absence painstakingly.

There also seems to be no way around positing discontinuous constituents (McCawley 1998), pace Evans and Levinson (2009), although the fieldworker should posit them only when no other analysis is reasonable.

Some constituents, such as sentence fragments, cannot be readily analyzed or described with the sentence, but belong to the realm of discourse and conversational analysis. They may pose a problem for the fieldworker, who will not be able to determine immediately what are full sentences, clauses, or phrases, and what are fragments. The fieldworker will have to rely on native speaker judgments to help recognize and distinguish fragments from grammatical clauses. Sadock and Zwicky (1985:187–191) distinguish the following sorts of sentence fragments:

- Free constituents
- Answers to questions, which themselves can be subdivided into
  - Question particle answers
  - Yes/no answer systems
  - Agree/disagree answer systems
  - Echo answer systems

Other useful terms having to do with constituency are: the (syntactic) phrase, a constituent that has one head and a dependent (see Section 11.4.5); the modifier, a dependent constituent that can be present in a phrase, but that is not required by the phrase; and the complement, a dependent constituent that is required to be in a particular phrase. It is important not to confuse the term complement as defined above with complement as used in the term complement clause, discussed in Section 11.4.5, or in the term complement, illustrated under "Noun Phrases and Verb Phrases".

The constituent called specifier is used in X-bar theory, and is another sort of dependent constituent. It occurs immediately under the XP, dominating the head constituent and complement construction. Specifiers such as auxiliary verbs and determiners in English seem to us to be theory-internal (in the sense used in this chapter) and suspiciously far easier to discover in Indo-European languages than in others. The difference between modifiers and specifiers is also discussed in Johnson (1992:122–136). The fieldworker will find the distinction between modifier and complement useful, but we doubt s/he will find the term specifier useful. For phrase structure grammar and X-bar theory, Radford (1988) is accessible and thorough.

In the rest of this section, we discuss the concept of the clause, and the most common sorts of phrases found in clauses, i.e. Noun Phrases, Verb Phrases, and Verbless predicate phrases.

11.4.4.1 Clauses

A clause can be defined as a syntactic constituent that contains one independent and finite verb.[14] If a clause is itself an independent constituent it is indistinguishable from a sentence. All languages appear to have a constituent identifiable as a clause, but it is not at all clear whether all languages have a clear distinction between sentences and clauses. The fieldworker will therefore spend more time in identifying clauses than in worrying whether the language has sentences. Further discussion of sentences is in Section 11.4.5.

Clause types can be classified by the sorts of constituents they contain, in particular by the sorts of verb phrases or predicates they contain. A very useful account of clause types is in Dryer (2007b).

11.4.4.2 Noun Phrases

The noun phrase (NP) is a universal constituent type (Bickford 1998:72). Within the NP there can be: Quantifier Phrases, which can contain Degree Phrases, and Adjective Phrases, which can also contain Degree Phrases. The NP can also have constituents traditionally called Prepositional or Postpositional Phrases, and Determiner Phrases, which are problematic, and are discussed in Section 11.4.5. Surveys of noun phrases are in Rijkhoff (2002) and in Dryer (2007c).

NPs often contain nominalized deverbal material. See Comrie and Thompson (1985:370–384, 2007:355–368) for a useful survey of the syntax of lexical nominalization: the nominalized material can assimilate to NP syntax; retain sentence syntax; or only partially assimilate, i.e. have characteristics of both NP syntax and sentence syntax. Nominalized complements that are NP constituents are further discussed in Noonan (1985:60–62, 2007). An example of a nominalized complement, formally an NP functioning as a subject, in brackets, is [*Algernon's shooting of the aardvark*] *drew international attention* (Noonan 1985:60).

---

[14] The term finite is not extremely useful cross-linguistically, and is useful only in families such as Indo-European. Finite verbs are typically inflected for subject and/or for TAM, whereas nonfinite verbs seem to be derived nominal, adjectival or adverbial forms, or morphologically subordinate in some way. Most languages of the world probably do not have anything like the difference between finite verbs and non-finite verbs.

11.4.4.3 Verb Phrases or Verbal Predicates

Another important constituent is the verb phrase (VP), which, depending on the theory, can be defined in many ways. Tests for identifying VPs have tended to be more theory-internal than those for NPs. Some theories predict that a VP is a universal constituent type, whereas others will predict that a language need not have an identifiable VP. There are special problems of course, in identifying a VP in nonconfigurational languages and in languages with a Verb–Subject–Object or an Object–Subject–Verb order (Kiss 1995) A useful approach is to define the VP very broadly as a predicate (i.e. everything but the subject NP) that has a verb as its head.

Predicates can include all sorts of complements. The most detailed survey of complement taking predicates is Noonan (1985:110–132), and his terminology is recommended here. Noonan distinguishes the following (one of Noonan's English examples is given after each term in parentheses, with the predicate in brackets):

- Utterance predicates (*Zeke [said that Norm left.]*)
- Propositional attitude predicates (*It [is certain that Hugh will be defeated.]*)
- Pretence predicates (*I [pretended that Ivan came.]*)
- Commentative or factive predicates (*Nelson [regrets that Perry got the nod.]*)
- Predicates of knowledge or acquisition of knowledge (*I [saw that Floyd left.]*)
- Predicates of fearing (*He [is afraid that Floyd came.]*)
- Desiderative predicates (*I [hope that John came.]*)
- Manipulative predicates (causative and permissive) (*Max [persuaded Nellie to run for mayor.]*)
- Modal predicates (*Vladimir [can eat a whole pizza.]*)
- Achievement or implicative predicates (*Zeke [tried eating spinach.]*)
- Phasal or aspectual predicates (*He [continued going down.]*)
- Immediate perception predicates (*The woman [saw the man stealing the chicken.]*)
- Negative predicates (not found in English; less common; a negative element functions as a verb-taking a predicate)
- Conjunctive predicates (not found in English; less common; a conjoining element (e.g. "and then") functions as a verb-taking a predicate

The first type, Utterance predicate, of course contains indirect or reported speech. Reported speech tends to have its own syntactic characteristics. Studies of reported speech are in Janssen and van der Wurff (1996).

A survey of intransitive predicates is in Stassen (1997), and a survey of adjectival predicates is in Wetzer (1996).

For comparison with English, see also the useful discussion of predicate types and complementation in Quirk et al. (1985:1150–220), and the comprehensive discussion of verb classes in Levin (1993).

11.4.4.4 Verbless Predicate Phrases

The fieldworker should also be aware of the interesting case of predicates without verbs. These tend to be simpler in structure than verb phrases, but they are often quite prominent in text frequency. Types of predicates without verbs include many of the structures called nominal sentences in traditional Indo-Europeanist and Semiticist literature. A survey of non-verbal predication is Hengeveld (1993). A useful typology is in Payne (1997:111–126), who distinguishes the following types of clauses without verbal predicates:

- Proper inclusion clauses ("s/he teacher", meaning 'S/he is a teacher.') (predicate nominal)
- Equative clauses or identification clauses ("He my father", meaning 'He is my father.') (predicate nominal)
- Attribution or characterization clauses ("John tall", meaning 'John is tall.' (predicate adjective)
- Location clauses ("The book on the table", meaning 'The book is on the table.') (predicate locative)
- Existential or presentative clauses ("There a bee on your apple", meaning 'There is a bee on your apple.')
- Possessive clauses ("Three cats to Mary" or "Mary's cats three", meaning 'Mary has three cats.')

As reflected in the English translations, the prominence of the verb "to be" is typical of most Indo-European languages.

A structural classification of clauses with non-verbal predicates or with semantically weak lexical verbal predicates is as follows. This classification is also from Payne (1997).

- NP + NP juxtaposition, or if adjectives exist.
- NP + Adj juxtaposition, not necessarily in that order. Sometimes in simple NP + NP juxtaposition some sort of declarative, topic, or evidential marker becomes obligatory on the predicate, as in Quechua.
- NP + Copula + NP, not necessarily in that order. The copula can formally be a verb, a pronoun, an invariant word (often called a particle), or a derivational morphological element. If the copula is a verb, it is often present in non-present tenses only, as in Russian or Quechua.
- NP + Semantically empty verb "to be" or "to have"+NP, not necessarily in that order. The semantically empty verb can be a denominal verbalizing affix in polysynthetic languages such as Eskimo.

A survey of copulas is in Pustet (2003), and a discussion of copula clauses and verbless clauses is in Dixon (2010b:159–188).

11.4.5 Headedness and Dependency

As mentioned in Section 11.2.2, many morphological or syntactic constructions can be seen as composed of a head constituent, and a dependent constituent. However, there are other constructions, both morphological and syntactic, that do not appear to have heads, and by implication they do not have dependents, either. It is a mistake for the fieldworker to try too hard to find a head and a dependent in every construction.

As seen in Section 11.3.3.3, headlessness is quite common in morphology, even in something as simple as compounding. Syntactic constructions tend to be headed, but here too there are exceptions. In the idiom *She kicked the bucket* meaning ‘she died’, it is difficult to argue that *kick* is the head of the verb phrase in the same sense that it is in the non-idiomatic reading of *kick the bucket*.

In formal syntax, terms such as Determiner Phrase (DP) and Prepositional Phrase (PP) are used. We assume from this usage that the determiner *the* is the head of the phrase in the DP *the dog*, and that the Preposition *in* is the head of the phrase in the PP *in school*, but following our definition of a head, and like Dixon (2010a:231) we find this far from obvious. Both in preparatory reading and in language description, the fieldworker needs to exercise caution with such dubious concepts, to make sure that his/her grammars will still be easy to read 300 years from now. The same is true of course with things such as Tense Phrases, Infl Phrases, Topic Phrases, Focus Phrases, Agreement Phrases, and so on.

The relationship between a head and its dependent is called dependency. The term Valence (or Valency) is used to refer to the dependency relationship between a verb and its argument phrases. See also “Valence or Valency” under Section “Grammatical Categories Characteristic of Verb Phrases”. The term valence is, however, somewhat tricky, because it can be used from a syntactic point of view or from a semantic point of view. The syntactic side of valence is called subcategorization – basically the number of complements which must be included under the VP with that verb. Headedness and dependency are relevant only to the syntactic side of valence. The semantic side of valence is called selectional restriction – basically the number of semantic roles (Section “Grammatical Categories Characteristic of Noun Phrases”) that go with the verb; this semantic usage has nothing to do with heads and dependents (Bickford 1998:75–92, Payne 1997:169–170, Whaley 1997:183–185, Kroeger 2004:11). Other constructions having to do with verbal dependency are serial verb constructions, and converbs or non-finite verb constructions, to which we now turn.

Serial verb constructions are typically phrases or clauses containing a head verb and a dependent verb, and the two verbs share arguments. Serial verb constructions have been discussed in Lord (1993), Alsina et al. (2001), Kroeger (2004:222–256), and Aikhenvald and Dixon (2006); summaries are in Payne (1997:307–312) and Whaley (1997:274–276). Great caution needs to be exercised in describing potential serial verb constructions, since it is not always clear that they are headed, or whether they are phrases or clauses, cases of clause chaining or converbs (see below), or even morphological compounds.

The term converb is often applied to verbs with potentially subordinating morphology in Altaic, South Asian, and more recently in African languages (Amha and Dimmendaal 2006). Other terms for non-finite verb forms such as gerunds, participles, and infinitives are primarily applied to Indo-European languages (Nedjalkov 1998). Therefore all of these terms should be used and defined with caution (Haspelmath and König 1995; König 1995).

Headedness and dependency are particularly relevant to the combining of phrases and clauses. Coordinate constructions are not headed, regardless of whether they are combinations of two phrases or combinations of two clauses; subordinate constructions are always headed.

The term Subordination (or Hypotaxis) are used to refer to the dependency between two clauses. This is a feature generally assumed to exist in all languages, though see Everett (2005:628–631) for potential counterevidence. The head clause is hardly ever referred to as a "head"; it is generally called the independent (or potentially independent) clause, superordinate clause, matrix clause, higher clause, or upstairs clause. The dependent clause is also called a subordinate clause, embedded clause, lower clause, or downstairs clause (Bickford 1998:328). Useful discussions of subordination are in Thompson and Longacre (1985), Payne (1997:317–320), Whaley (1997:247–280), Bickford (1998:328), and Thompson et al. (2007). The most detailed typological survey of subordination is Cristofaro (2003).

It is important to distinguish between Subordination and Complementation, although there is wide overlap between these two concepts. Since, as Dixon (2010b:370) puts it: "a complement clause is a clause which fills an argument slot in the structure of another clause", there is no doubt that all complement clauses are subordinate clauses, although of course not all subordinate clauses are complement clauses. In addition, not all complements are clausal, and other mechanisms for expressing complementation such as serial verb constructions, apposition, or clause chaining may be used. Dixon (1995) introduced the useful distinction between complement clauses and these other mechanisms, which he called complementation strategies. A good survey of complement clauses and complementation strategies is in Dixon (2010b:370–412), and a collection of articles on complementation written within Dixon's framework is Dixon and Aikhenvald (2006).

Functional accounts of subordination and complementation are Ransom (1986) and Haiman and Thompson (1988), and interesting more formally oriented accounts are in de Geest and Putseys (1984).

Foley and Van Valin (1984:240–263), following Olson (1981), distinguish between embedded dependent clauses, which they call subordinate clauses, and non-embedded dependent clauses, which they call cosubordinate clauses. The distinction they draw between subordination and cosubordination is not likely relevant for very many languages, although it is adopted by Whaley (1997:267–280).

Adjectival subordinate clauses, generally called relative clauses, can be externally headed, internally headed, or headless. Note that "headless" is a somewhat confusing term. To be sure, since all relative clauses are dependent clauses, they have heads by definition. Headless relative clauses can be viewed as relative clauses in which the head is not present, but is identifiable from context. The fieldworker needs to be aware that not everything translated as a relative clause is analyzable as a relative clause; some languages use adverbial subordinate clauses, nominalizations, complement clauses, or even coordination to express something that is translated into English as a relative clause (Hale 1976).

For surveys of relative clauses, we refer to Lehmann (1984), Keenan (1985b), Payne (1997:325–336), Whaley (1997:259–265), Comrie (1998), Bickford (1998:345), Kroeger (2004:165–191), and Andrews (2007b). The most recent survey is the very well written and commonsensical chapter in Dixon (2010b:313–369).

Adverbial subordinate clauses are generally easier to analyze than relative clauses. Surveys of adverbial clauses are in Thompson and Longacre (1985), Kortmann (1997), Payne (1997:316–320), Whaley (1997:250–255), and Thompson et al. (2007).

Many languages exhibit clause chaining, a phenomenon characterized by long strings of dependent clauses, all loosely dependent on one final independent clause (Payne 1997:321; Kroeger 2004:242–250). Longacre (1985, 2007) is a fairly idiosyncratic discussion of clause chaining, with examples from the languages of Meso-America, Papua New Guinea, and the Philippines.

Although in many cases the difference between dependent and independent clauses is clear, the morphology involved in complementation, subordination, and clause chaining does not necessarily help the fieldworker make syntactic dependency distinctions. In such cases, it becomes difficult to determine what constitutes the boundaries of the sentence, or if "sentence" is even a relevant term for the language in question. Indeed, as sentences become very long, it is less likely that they can be analyzed in terms of dependent and independent clauses, and it is more likely that they can be analyzed in terms of clauses and paragraphs. Thus when translating texts, the fieldworker will often encounter the "endless sentence" (Longacre 1985:282–283). When working with naturally occurring discourse, the fieldworker should begin with the assumption that speakers cannot identify clause boundaries, especially at the level of the combined clause. The problem is especially acute with subordinate clauses, which can be mistakenly represented as independent clauses. On the other hand, speakers of clause-chaining languages do not have problems identifying finite verbs at the end of clause chains.

One last point concerning subordination is that the historical origins of subordinating morphemes or words can often be traced back to other elements. One typical case is the development of subordinators from postpositions in Tibeto-Burman languages (Genetti 1986, 1991).

There is one case where the term dependency has nothing to do with headedness; these are the headless constructions we will call mutual dependencies. Mutual dependencies across syntactic elements – which can potentially be even a clause apart from each other – have not been discussed very much in the recent syntactic literature, even though they were well-known to the Neo-grammarians. Payne (1997:336) discusses a type of mutual dependency called correlative structures. Examples of such structures are:

- Adverbial structures, such as "when…, then…" or "where…, there…"
- Conditional structures, such as "if…., then…"
- Causal structures, such as "because …., or that is why…"
- Relative structures, such as "whoever…, s/he…," or "the one who…, that one…"

Like parataxis, correlative structures are more common in spoken than in written discourse.

Another type of mutual dependency is the framing device. Many languages, including Western Apache, have ways of framing quotation of direct speech.

A possible frame is: "s/he said: [material directly quoted], thus s/he said". The structure, then, functions as "syntactic quotation marks". Frames are also common with focusing devices, such as negative marking (French, Middle Dutch, Navajo), in deictic elements, such as demonstratives in French or in Mayan languages, and in question particles (Western Apache).

In some Germanic languages such as German and Dutch, a finite verb is the second constituent of a sentence, and there is often a nonfinite verb at the end of the same clause. These could also be considered verbal frames.

11.4.6 Modularity

In the previous sections, we have assumed that syntax is basically a higher or broader continuation of morphological constructions, or alternatively, that morphology is a lower and more fine-grained continuation of syntactic constructions. To a large extent, this view is valid, but different theories are coming to regard grammar, and in particular syntax, as modular. The idea is that there is not one hierarchical structure building from phoneme to discourse, but rather that there a number of independent modules, each of which can, to some extent, impose its own hierarchical structure on a surface structure sequence of words. Certain sorts of constructions will then be predicted to arise from mismatches between two or more modules, segmenting the same sequence of words in different ways. Our favorite modular view of grammar is Sadocks' (1991) model of Autolexical syntax, now called automodular. Sadock posits, for example, a morphological module and a syntactic module; these come together to explain constructions that have been problematic for morphology and for syntax – such as cliticization and noun-incorporation – but that can be elegantly accounted for by postulating an interface between a syntactic and a morphological module.

There are several unresolved questions about the model which the fieldworker will need to consider. One is: how many modules are there? The other is: which phenomenon corresponds best with which module? Our view is that cliticization is perhaps best explained not by appealing to an interface between morphological and syntactic modules, but to an interface between a phonological word module and a syntactic word module. Our advice to the fieldworker is to carefully justify positing any new module, and to avoid the temptation to explain syntactic phenomena in modular terms where a non-modular account will do.

11.4.7 Syntactic Mechanisms

There are various formal mechanisms in syntax. Several have been described already in the context of morphosyntactic marking in Section 11.2.2. Here we discuss argument alignment systems and reference-tracking mechanisms.

Argument systems (or case alignment systems) group together – through syntax and/or morphology – the core argument of an intransitive verb (abbreviated as S), the agentive argument of a transitive verb (abbreviated as A), and the patient-like argument of a transitive verb (abbreviated as O). Basic references are Sapir (1917), Andrews (1985, 2007a), Dixon (1994), Mithun (1991), and Malchukov and Spencer (2008). Summaries are in Payne (1997:133–135), Whaley (1997:151–169), and Kroeger (2004:280–82). The three basic argument alignment systems are:

- Nominative/accusative (or accusative, for short): S and A are treated alike and called nominative, and O is called accusative.
- Ergative/absolutive (or ergative, for short): S and O are treated alike and called absolutive, and A is called ergative.
- Split intransitive, split stative, or split S: For some intransitive verbs, the core argument is marked like the O of a transitive clause, and for other intransitive verbs, the core argument is marked like the A of a transitive clause, depending on a variety of syntactic or semantic parameters.

A tripartite system, which treats A, O and S separately, is possible, and is attested in Wangkumara of Australia, but it is cross-linguistically quite rare (Whaley 1997:158).

In accusative systems like that of the Andean language Quechua, the nominative tends to be zero while the accusative is marked, but there are plenty of counterexamples. For instance, nominative and accusative were both marked in the older Indo-European languages, such as Latin, Greek, or Sanskrit.

In ergative systems, the ergative tends to be marked and the absolutive tends to be zero, and there are far fewer exceptions. Systems called split ergative typically exhibit accusative alignment in present or imperfective verbs, and ergative alignment in past or perfective verbs (Kazenin 1994; Whaley 1997:162). The problems of syntactic ergativity, a theory-internal notion, are discussed in Payne (1997:162–166) and Kroeger (2004:280–309).

Accusative systems tend to have a passive construction, while ergative systems tend to have an antipassive construction.

It is not necessarily true that the three alignment systems outlined above are the basic ones from which all case marking systems can be derived. For example, in some languages, subject or object marking seems to be determined by the definiteness of the NP (de Hoop and de Swart 2008). Other pragmatic, semantic, or discourse-related factors that determine case marking are discussed in Barðdal and Chelliah (2009). See Aikhenvald et al. (2001) for additional discussion of non-canonical case marking.

Reference tracking mechanisms track arguments across clauses, and are surveyed by Foley and Van Valin (1984:321–366), and by Comrie (1989). Terms and oppositions that the fieldworker will find useful are: switch-reference systems (see Section 11.2.2), as opposed to switch-function systems[15]; coreference, as

[15] The term switch-function system never caught on in the way that switch-reference did, probably because it simply refers to the normal English way to maintain reference, as in the subject of passive of the third clause in the sentence *John went to work and talked to his boss and was given a promotion* (Foley and Van Valin (1984:354)

opposed to disjoint reference; pronominalization, as opposed to full NPs; anaphoric as opposed to cataphoric relationships; ellipsis; and antecedent reference. A typology of anaphora is in Huang (2000).

Reference tracking mechanisms track referents not only across two clauses within a single sentence, but also across sentences, and even across paragraphs; thus the fieldworker cannot expect to be able to separate, a priori, syntactic reference tracking mechanisms from discourse-based reference tracking mechanisms. Depending on one's theory, this might mean that discourse structure is really some sort of higher-level syntax, or it might mean that there is no distinction to be made between discourse and syntax, with syntax nothing but a grammaticized type of discourse. The fieldworker actually need not take a firm position in these debates, but instead must responsibly describe the whole system, wherever that description might lead.

For the related matters of topic or referent continuity, action continuity, and thematic continuity, a good overview is Payne (1997:342–353), and a quantitative functionalist study is Givón (1983).

11.4.8 Fuzziness, Clines, and Grammatical Hierarchies

The final syntactic issue that is useful for fieldworkers is the concept of fuzziness. One good question to ask about any grammatical or lexical category or syntactic notion is how fuzzy it is. The idea of fuzzy or gradient grammar goes back to Bolinger (1961a, b). A very useful reader with all the classic statements on fuzzy grammar is Aarts et al. (2004).

In the formalist frameworks, fuzziness was not a helpful idea, and the structural idea of all-or-none, or plus or minus features won the day, as one can see in the [±V] [±N] usages to define lexical categories, and [±anaphor] [±pronominal] taxonomies, and Pro-Drop or Null subject parameters in Chomsky's writings (e.g. Chomsky 1981). Other such all-or-none parameters devised by formalists are the Pronominal argument parameter (Jelinek 1984), and the Polysynthesis parameter (Baker 1996).

Notwithstanding the appeal of such crisp neatness, the fieldworker will find that such usages and parameters are theory-internal in that they depend heavily on syntactic tests which are difficult to perform. It is not uncommon, for example, for formalist A to decide, on the basis of fieldwork, that language X is a Pronominal argument language, but for formalist fieldworker B to decide that language X is actually not a Pronominal argument language. Has progress been made? Has formalist B improved upon formalist A's analysis? Maybe, maybe not!

Studies in prototype theory (Rosch 1975) and interest in cognitive grammar (Langacker 1991; Geeraerts and Cuyckens 2007) have popularized the idea of prototypical constructions with fuzzy edges. We do agree, however, that some things are not fuzzy. Thus we think that transitivity is not a scalar matter, *pace* Hopper and Thompson (1980), but it is obvious that this is a terminological matter, and we wish that Hopper and Thompson had given their scale another name.

One domain of syntax where prototypicality studies have blossomed is the area of typological syntactic universals. The idea is that typological universals are not categorical but implicational, and arranged on hierarchical clines or scales. Let us assume, for example, that construction A is the prototypical one, construction B is less prototypical, and construction C even less so, etc. The implicational universal would then say that every language that has C also has A and B, and every language that has B also has A.

Here is a list of the major hierarchies and scales relevant to syntax, with references:

- Animacy hierarchy (Silverstein 1976; Dixon 1979; Payne 1997:150; Croft 2003:128–132)
- Binding hierarchy (Givón 1980; Croft 2003:214–216)
- Definiteness hierarchy (Comrie 1981; Payne 1997:150; Croft 2003:132)
- Dependency between grammatical system hierarchies (Aikhenvald and Dixon 1998)
- Grammatical integration scale (Payne 1997:306)
- Grammatical relation hierarchies, such as the hierarchy of NP accessibility to relativization (Keenan and Comrie 1977; Fox 1987; Payne 1997:335; Croft 2003:142–154; Kroeger 2004:93–94)
- Number hierarchy (Corbett 2000; Croft 2003:126–128)
- Scale of Predicate types according to their likelihood of lacking a lexical verb (Payne 1997:113)
- Prepositional noun modifier hierarchy (Hawkins 1983; Croft 2003:122–123)
- Subject construction hierarchy (Croft 2001, 2003:199–200)
- Subordinate clause deranking hierarchy (Cristofaro 2003; Croft 2003:217–219)

A recent position in cognitive linguistics holds that language universals do not exist; a strong argument for this position is in Evans and Levinson (2009). It is true that the arguments for the universal existence of a word-class ‘adjective’ are weak; and the arguments for the universal existence of a noun–verb distinction are weak as well. However, we consider denying the existence of constituency – on the basis of free word order phenomena as in Latin poetry or some Australian languages – to go too far. Surely, in a language such as Latin where parts of constituents can be separated from each other by other material, speakers know what goes with what in terms of dependencies. Denying the existence of discontinuous constituency seems to us to be overly dogmatic about what constituency means. Maybe the best advice to the fieldworker on this issue is to stay away from arguments for or against universals, and to stick with describing the language with whatever categories and mechanisms s/he needs. After all, the search for universals should not be a fieldworker’s primary concern.

Another recent trend in cognitive linguistics is the study of linguistic complexity (Gil 2009; Sampson 2009). This is relevant to issues of universals because field linguists have generally held the view that there is one universally true fact about languages: they are all just about equally complex. However, with more naturalistic data coming in from many languages, we might question whether even that idea is a universal. As for the issue of universals, it is not necessary for field linguists to take a firm position on the issue of complexity: if the grammar of a language appears to be simple, they should describe it using simple terms, categories, and constructions; if the grammar of a language appears to be complex, they should describe they should describe it using complicated terms, categories, and constructions.

Furthermore, linguistic complexity cannot necessarily be related to the unconscious language competence of the speaker, and linguistic complexity may not necessarily – as Gil (2009) and Sampson (2009) have suggested – be the result of different evolutionary paths. There is evidence that, in some parts of the world (such as Australia and Papua New Guinea), speakers in small scale societies put so much value on linguistic diversity that when two groups separate, conscious efforts are made to make the languages more complicated in different ways, so that special language learning efforts have to be made if a speaker wants to belong to a particular community. This conscious production of linguistic diversity and obscurity has been called "esoterogeny" (Evans 2010:12–14). Again, the fieldworker has to be aware of this sort of phenomenon, which is quite unknown in large scale societies.

11.5 Conclusions

The reference section which concludes this chapter is intended as a beginning resource for fieldworkers interested in a specific topic. Obviously, no one is going to read all of these works before leaving for the field. It is important, however, to decide what needs to be read before going to the field and what needs to be consulted while in the field.

Before the fieldtrip, it is important to decide which of the reference works the fieldworker will take, or what pages out of these works s/he would like to have copies of. Not everything is on the Internet, and the fieldworker might not have Internet access in the field. S/he can only take a few books and articles, and it is hoped that this chapter, intended as a quick and dirty survey, will help the fieldworker decide what to read in the field, what to read in between trips, and what to read after fieldtrips. Chapter 12 provides a practical guide to eliciting information related to the issues in this chapter.

Let us finish with some comments on terminology. Fieldworkers (and linguists in general, as pointed out by Cristofaro 2006:157–161) need to use a standard terminology for linguistic categories and structures. Adopting a cavalier attitude towards terminology by using terms in nonstandard ways, or by coining new terms for constructions that are well known, will impede cross-linguistic comparison and can even detract from the pleasure a grammar enthusiast looks for in reading a grammar. If we look to the hard sciences such as mathematics, physics, and chemistry, we find very little tolerance for terminological idiosyncrasy. It is true that every grammatical tradition and specialists in every language family have their own specialized terminology, and these traditions are often difficult to change. The fieldworker describing a language for the first time has the additional responsibility of getting the terminology right within his or her tradition. In addition to using terms found in the typological surveys mentioned in Section 11.2.1.1, we suggest checking the developing General Ontology for Linguistic Description (GOLD)[16] and the Leipzig Glossing Rules,[17] both of which offer a useful standard set of glossing abbreviations. The chapter on Terminology in Dixon (2010a:214–241) and the sections called "Avoid sloppy terminology" (Dixon 2010a:75–80) are also well worth reading.

The following is a set of commonly found terms to be avoided or to be used only after they have been clearly defined, as they are either too vague or ambiguous to be useful, or are used inconsistently in existing grammars.

- *Clitic*. This term should only be used when cliticization is involved. There is a tendency in recent formalist literature to use this for certain types of affixes. See also *Particle* below.
- *Complex.* Not every complicated construction should be called complex. The term "complex construction" should be reserved for constructions involving embedding or subordination. The term "complex predicates" or "complex predication" (Seibert 2000) should be clearly defined, as it can refer to a host of constructions including: verb compounding, serial verb constructions, noun-incorporation, and adverbial modification of verbs.
- *Govern and Government.* We agree with Dixon (2010a:231) that this term of traditional grammar is used in confusing ways and can be done without.
- *Emphatic*. This term can be used to refer to stress, insistence, topic, focus, or intensifiers. Therefore, it is too vague a term. If a particular morpheme or construction is called "emphatic", chances are that the fieldworker has not discovered what its precise meaning is.
- *Marker*. It is preferable to use terms that are more specific as to the form and function of the "marker". For example, subject or object "markers" can formally be particles, clitics, or affixes; functionally, they can indicate case or agreement, or they might just be pronominal elements (Bickford 1998: 270).
- *Marking*. As we discussed in Section 11.2.2, when using this term, it is absolutely necessary to distinguish between formal marking and functional or semantic marking. We also recommend Dixon (2010a:235–240) regarding the concept of markedness, which should not be confused with marking.
- *Particle*. This term can designate any class of short uninflectable word. It can be used as a cover term, but in specific cases of grammatical description, it is usually too vague and a more specific term should be decided upon instead. Also, it is important not to confuse *clitic,* which is a formal term (discussed in Section 11.3.3.1), with *particle*, which refers to a lexical class. A particle can be a clitic, and a clitic can be a particle, but a particle is not the same as a clitic.

---

[16] http://linguistics-ontology.org/ns/gold/0.45/treeview/treeview.html
[17] http://www.eva.mpg.de/lingua/pdf/LGR08_09_12.pdf

- New terms referring to voices or cases that start with *anti-* or *un-*. We find such terms to be uninformative and unilluminating. It seems that we are stuck with "unaccusative" and "antipassive", but do we really need "unergative", "anticausative", "antiergative" (LaPolla 1992), or "antidative" (Dryer 1986)? We tend to agree with Dixon's (2010b:155–156) comments on confusing uses of the terms "unaccusative" and "unergative".

References

Aarts, Bas, David Denison, Evelien Keizer and Gergana Popova, eds. 2004. *Fuzzy Grammar. A Reader.* Oxford: Oxford University Press.

Abbi, Anvita. 1992. *Reduplication in South Asian Languages. An Areal, Typological and Historical Study*. New Delhi: Allied Publishers.

Abbi, Anvita. 2001. *A Manual of Linguistic Field Work and Indian Language Structures.* (Lincom Handbooks in Linguistics 17.) Munich: Lincom Europa.

Abraham, Werner. 1991a. Modal particle research: the state of the art. *Multilingua* 10(1/2):9–15.

Abraham, Werner, ed. 1991b. *Discourse Particles: Descriptive and Theoretical Investigations in the Logical, Syntactic and Pragmatic Properties of Discourse Particles in German*. Amsterdam and Philadelphia: John Benjamins.

Aikhenvald, Alexandra Y. 2000. *Classifiers: A Typology of Noun Categorization Devices*. (Oxford Studies in Typology and Linguistic Theory) Oxford: Oxford University Press.

Aikhenvald, Alexandra Y. 2004a. Nominal classification: towards a comprehensive typology. *Sprachtypologie und Universalienforschung* (Focus on Nominal Classification, ed. by Alexandra Aikhenvald), 57(2/3):105–116.

Aikhenvald, Alexandra Y. 2004b. *Evidentiality*. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y. 2007. Typological Distinctions in Word-Formation. In *Language Typology and Syntactic Description. Vol. 3: Grammatical categories and the lexicon*, ed. by Timothy Shopen, 1–65. Cambridge: Cambridge University Press. [2nd edition]

Aikhenvald, Alexandra Y., and R. M. W. Dixon. 1998. Dependencies between grammatical systems. *Language* 74(1):56–80.

Aikhenvald, Alexandra Y., and R. M. W. Dixon, eds. 2003. *Studies in Evidentiality*. (Typological Studies in language 54.) Amsterdam and Philadelphia: John Benjamins.

Aikhenvald, Alexandra Y., and R. M. W. Dixon, eds. 2006. *Serial Verb Constructions. A Cross-Linguistic Typology*. (Explorations in Linguistic Typology 2.) Oxford: Oxford University Press.

Aikhenvald, Alexandra Y., R. M. W. Dixon, and Masayuki Onishi, eds. 2001. *Non-Canonical marking of Subjects and Objects*. (Typological Studies in Language 46.) Amsterdam and Philadelphia: John Benjamins.

Allen, J. H., and James B, Greenough. 1931. *New Latin Grammar for Schools and Colleges*. Boston, MA.: Ginn & Company.

Alsina, Alex, Joan Bresnan, and Peter Sells, eds. 2001. *Complex Predicates*. Stanford: CSLI.

Amha, Azeb, and Gerrit Dimmendaal. 2006. Converbs in an African perspective. In *Catching Language: The Standing Challenge of Grammar Writing* ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 393–440. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Anderson, Gregory D. S. 2005. *Auxiliary Verb Constructions*. (Oxford Studies in Typology and Linguistic Theory.) Oxford: Oxford University Press.

Anderson, Stephen R. 1982. Where's morphology? *Linguistic Inquiry* 13:571–612.

Anderson, Stephen R. 1985a. Typological distinctions in word formation. In *Language Typology and Syntactic Description. Vol. 3: Grammatical categories and the lexicon*, ed. by Timothy Shopen, 3–56. Cambridge: Cambridge University Press.

Anderson, Stephen R. 1985b. Inflectional morphology. In *Language Typology and Syntactic Description. Vol. 3: Grammatical categories and the lexicon*, ed. by Timothy Shopen, 150–201. Cambridge: Cambridge University Press.

Anderson, Stephen R. 1992. *Amorphous Morphology*. Cambridge: Cambridge University Press.

Anderson, Stephen R., and Edward L. Keenan. 1985. Deixis. In *Language Typology and Syntactic Description. Vol. 3: Grammatical Categories and the Lexicon*, ed. by Timothy Shopen, 259–308. Cambridge: Cambridge University Press.

Andrews, Avery D. 1985. The major functions of the noun phrase. In *Language Typology and Syntactic Description. Vol. 1: Clause Structure*, ed. by Timothy Shopen, 62–154. Cambridge: Cambridge University Press.

Andrews, Avery D. 2007a. The major functions of the noun phrase. In *Language Typology and Syntactic Description. Vol. 1: Clause Structure*, ed. by Timothy Shopen, 132–223. Cambridge: Cambridge University Press. [2nd edition].

Andrews, Avery D. 2007b. Relative clauses. In *Language Typology and Syntactic Description. Vol. 2: Complex Constructions*, ed. by Timothy Shopen, 206–236. Cambridge: Cambridge University Press. [2nd edition]

Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. (Linguistic Inquiry Monograph 1.) Cambridge, MA: MIT Press.

Aronoff, Mark. 1994. *Morphology by Itself: Stems and Inflectional Classes*. (Linguistic Inquiry Monograph 22.) Cambridge, MA: MIT Press.

van der Auwera, Johan, and Donall P. O'Baoill, eds. 1998. *Adverbial Constructions in the Languages of Europe*. Berlin: Mouton de Gruyter.

Baker, Mark. 1996. *The Polysynthesis Parameter*. Oxford: Oxford University Press.

Barðdal, Jóhanna, and Shobhana L. Chelliah, eds. 2009. *The Role of Semantic, Pragmatic and Discourse Factors in the Development of Case*. Amsterdam and Philadelphia: John Benjamins.

Baron, Irene, Michael Herslund, and Finn Sorensen, eds. 2001. *Dimensions of Possession*. (Typological Studies in Language 47.) Amsterdam and Philadelphia: John Benjamins.

Bauer, Laurie. 2003. *Introducing Linguistic Morphology*. 2nd Edition. Washington, D.C.: Georgetown University Press.

Bauer, Laurie. 2004. *A Glossary of Morphology*. Washington D.C.: Georgetown University Press.

Bauer, Winifred, with William Parker and Te Kareongawai Evans. 1993. *Maori*. (Descriptive Grammar Series.) London and New York: Routledge.

Bedell, George. 1997. Agreement in Lai. In *SEALS V, Papers from the Fifth Annual Meeting of the Southeast Asian Linguistics Society*, ed. by Shobhana L. Chelliah and Willem J. de Reuse, 21–32. Tempe, AZ: Program for Southeast Asia Studies, Arizona State University.

Bickel, Balthasar. 1991. *Typologische Grundlagen der Satzverkettung. Ein Beitrag zur allgemeinen Grammatik der Satzverbindung und des Fährtenlegens*. (Arbeiten des Seminars für Allgemeine Sprachwissenschaft No. 9.) Zurich: University of Zurich.

Bickel, Balthasar, and Johanna Nichols. 2007. Inflectional morphology. In *Language Typology and Syntactic Description. Vol. 3: Grammatical Categories and the Lexicon*, ed. by Timothy Shopen, 169–240. Cambridge: Cambridge University Press. [2nd edition].

Bickford, J. Albert. 1998. *Tools for Analyzing the World's Languages: Morphology and Syntax*. Dallas, TX: Summer Institute of Linguistics.

Bird, Charles, and Timothy Shopen. 1979. Maninka. In *Languages and Their Speakers*, ed. by Timothy Shopen, 59–111. Cambridge, MA: Winthrop.

Blake, Barry J. 1994. *Case*. (Cambridge Textbooks in Linguistics.) Cambridge: Cambridge University Press

Boas, Franz. 1911. Introduction. In *Handbook of American Indian Languages*. Volume 1. (Smithsonian Institution. Bureau of American Ethnology. Bulletin 40), ed. by Franz Boas, 1–83. Washington, D.C.: Government Printing Office.

Bolinger, Dwight. 1961a. Syntactic blends and other matters. *Language* 37:366–81.

Bolinger, Dwight. 1961b. *Generality, Gradience, and the All-or-None*. The Hague: Mouton.
Bowern, Claire. 2008. *Linguistic Fieldwork. A Practical Guide*. New York: Palgrave MacMillan.
Brody, Jill. 1984. Some problems with the concept of basic word order. *Linguistics* 22:711–36.
Brown, Penelope, and Stephen C. Levinson. 1987. *Politeness. Some universals in language usage*. (Studies in Interactional Sociolinguistics 4.) Cambridge: Cambridge University Press.
Bybee, Joan L. 1985. *Morphology: A Study of the Relation between Meaning and Form*. (Typological Studies in Language 9.) Amsterdam and Philadelphia: John Benjamins.
Bybee, Joan L., Revere Perkins, and William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago, IL and London: University of Chicago Press.
Campe, Petra. 1994. *Case, Semantic Roles, and Grammatical Relations: A Comprehensive Bibliography*. Amsterdam and Philadelphia: John Benjamins.
Chafe, Wallace L., and Johanna Nichols, eds. 1986. *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, NJ: Ablex.
Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley, CA and Los Angeles, CA: University of California Press.
Chelliah, Shobhana L. 1997. *A Grammar of Meithei*. (Mouton Grammar Library 17.) Berlin and New York: Mouton de Gruyter.
Chelliah, Shobhana L. 2009. Semantic role to new information in Meithei. In *The Role of Semantics, Pragmatics, and Discourse in the Development of Case*, ed. by Jóhanna Barðdal and Shobhana L. Chelliah, 377–400. Amsterdam and Philadelphia: John Benjamins.
Chomsky, Noam A. 1957. *Syntactic Structures*. The Hague: Mouton.
Chomsky, Noam A. 1964. Current Issues in linguistic theory. In *The Structure of Language. Readings in the Philosophy of Language*, ed. by Jerry A. Fodor and Jerrold J. Katz, 50–118. Englewood Cliffs, N. J.: Prentice Hall.
Chomsky, Noam A. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
Chomsky, Noam A. 1981. *Lectures on Government and Binding*. Dordrecht, The Netherlands: Foris.
Chung, Sandra, and Alan Timberlake. 1985. Tense, aspect, and mood. In *Language Typology and Syntactic Description. Vol. 3: Grammatical Categories and the Lexicon*, ed. by Timothy Shopen, 202–258. Cambridge: Cambridge University Press.
Comrie, Bernard. 1976. *Aspect*. (Cambridge Textbooks in Linguistics.) Cambridge: Cambridge University Press.
Comrie, Bernard. 1981. *Language Universals and Linguistic Typology: Morphology and Syntax*. Oxford: Basil Blackwell.
Comrie, Bernard. 1985a. *Tense*. (Cambridge textbooks in Linguistics.) Cambridge: Cambridge University Press.
Comrie, Bernard. 1985b. Causative verb formation and other verb-deriving morphology. In *Language Typology and Syntactic Description. Vol. 3: Grammatical Categories and the Lexicon*, ed. by Timothy Shopen, 309–348. Cambridge: Cambridge University Press.
Comrie, Bernard. 1988. Linguistic typology. *Annual Review of Anthropology* 17:145–59.
Comrie, Bernard. 1989. Some general properties of reference-tracking systems. In *Essays on Grammatical Theory and Universal Grammar*, ed. by Doug Arnold, Martin Atkinson, Jacques Durand, Claire Grover, and Louisa Sadler, 36–51. Oxford: Clarendon Press.
Comrie, Bernard. 1998. Rethinking the typology of relative clauses. *Language Design* 1:59–86.
Comrie, Bernard, and Norval Smith. 1977. Lingua descriptive studies: Questionnaire. *Lingua* 42:1–72.
Comrie, Bernard, and Sandra A. Thompson. 1985. Lexical nominalization. In *Language Typology and Syntactic Description. Vol. 3: Grammatical Categories and the Lexicon*, ed. by Timothy Shopen, 349–398. Cambridge: Cambridge University Press.
Comrie, Bernard, and Sandra A. Thompson. 2007. Lexical nominalization. *Language Typology and Syntactic Description. Vol. 3: Grammatical Categories and the Lexicon*, ed. by Timothy Shopen, 334–381. Cambridge: Cambridge University Press. [2nd edition].

Corbett, Greville G. 1989. An approach to the description of gender systems. In *Essays on Grammatical Theory and Universal Grammar*, ed. by Doug Arnold, Martin Atkinson, Jacques Durand, Claire Grover, and Louisa Sadler, 53–89. Oxford: Clarendon Press.

Corbett, Greville G. 1991. *Gender*. (Cambridge Textbooks in Linguistics.) Cambridge: Cambridge University Press.

Corbett, Greville G. 2000. *Number*. (Cambridge Textbooks in Linguistics.) Cambridge: Cambridge University Press.

Corbett, Greville G. 2007. *Language Typology and Syntactic Description. Vol. 3: Grammatical Categories and the Lexicon*, ed. by Timothy Shopen, 241–279. Cambridge: Cambridge University Press. [2nd edition]

Corbett, Greville G. 2006. *Agreement*. (Cambridge Textbooks in Linguistics.) Cambridge: Cambridge University Press.

Craig, Colette G., ed. 1986. *Noun Classes and Categorization*. (Typological Studies in Language 7.) Amsterdam and Philadelphia: John Benjamins.

Creissels, Denis. 2006a. *Syntaxe générale. Une introduction typologique 1*. Catégories et constructions. Paris: Lavoisier.

Creissels, Denis. 2006b. *Syntaxe générale. Une introduction typologique 2*. La phrase. Paris: Lavoisier.

Cristofaro, Sonia. 2003. *Subordination*. (Oxford Studies in Typology and Linguistic Theory) Oxford: Oxford University Press.

Cristofaro, Sonia. 2006. The organization of reference grammars: A typologist user's point of view. In *Catching Language: The Standing Challenge of Grammar Writing*, ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 137–170. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Croft, William. 2001. *Radical Construction Grammar. Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

Croft, William. 2003. *Typology and Universals*. 2nd edition. (Cambridge Textbooks in Linguistics) Cambridge: Cambridge University Press.

Dahl, Östen. 1985. *Tense and Aspect systems*. New York: Basil Blackwell.

DeLancey, Scott. 1997. Mirativity: the grammatical marking of unexpected information. *Linguistic Typology* 1:33–52.

DeLancey, Scott. 2001. The mirative and evidentiality. *Journal of Pragmatics* 33:369–82.

Derbyshire, Desmond C. 1985. *Hixkariana and Linguistic Typology*. Dallas: SIL International.

Diessel, Holger. 1999. *Demonstratives: Form, Function and Grammaticalization*. (Typological Studies in Language 42.) Amsterdam and Philadelphia: John Benjamins.

Dixon, R. M. W. 1979. Ergativity. *Language* 55:59–138.

Dixon, R. M. W. 1986. Noun classes and noun classification in typological perspective. In *Noun Classes and Categorization*, ed. by Colette Craig, 105–112. Amsterdam and Philadelphia: John Benjamins.

Dixon, R. M. W. 1989. Subject and object in universal grammar. In *Essays on Grammatical Theory and Universal Grammar*, ed. by Doug Arnold, Martin Atkinson, Jacques Durand, Claire Grover, and Louisa Sadler, 91–118. Oxford: Clarendon Press.

Dixon, R. M. W. 1991. *A New Approach to English Grammar, on Semantic Principles*. Oxford: Clarendon Press.

Dixon, R. M. W. 1994. *Ergativity*. Cambridge: Cambridge University Press.

Dixon, R. M. W. 1995. Complement clauses and complementation strategies. In *Grammar and Meaning: Essays in honour of Sir John Lyons*, ed. by F. R. Palmer, 175–220. Cambridge: Cambridge University Press.

Dixon, R. M. W. 1997. *The Rise and Fall of Languages*. Cambridge: Cambridge University Press.

Dixon, R. M. W. 2003. Demonstratives: A cross-linguistic typology. *Studies in Language* 27(1):61–112.

Dixon, R. M. W. 2010a. *Basic Linguistic Theory. Volume 1 Methodology*. Oxford: Oxford University Press.

Dixon, R. M. W. 2010b. *Basic Linguistic Theory. Volume 2 Grammatical Topics*. Oxford: Oxford University Press.

Dixon, R. M. W., and Alexandra Y. Aikhenvald. 2000. *Changing Valency: Case Studies in Transitivity*. Cambridge: Cambridge University Press.

Dixon, R. M. W., and Alexandra Y. Aikhenvald, eds. 2002. *Word. A Cross-Linguistic Typology*. Cambridge: Cambridge University Press.

Dixon, R. M. W., and Alexandra Y. Aikhenvald, eds. 2004. *Adjective Classes. A Cross-Linguistic Typology*. (Explorations in Linguistic Typology 1.) Oxford: Oxford University Press.

Dixon, R. M. W., and Alexandra Y. Aikhenvald, eds. 2006. *Complementation. A Cross-Linguistic Typology*. (Explorations in Linguistic Typology 3.) Oxford: Oxford University Press.

Dixon, R. M. W., and Alexandra Y. Aikhenvald, eds. 2009. *The Semantics of Clause Linking. A Cross-Linguistic Typology*. (Explorations in Linguistic Typology 5.) Oxford: Oxford University Press.

Dooley Robert, A., and Stephen Levinsohn. 2001. *Analyzing Discourse: A Manual of Basic Concepts*. Dallas, TX: SIL International.

Downing, Pamela, and Michael Noonan, eds. 1995. *Word Order in Discourse*. (Typological Studies in Language 30.) Amsterdam and Philadelphia: John Benjamins.

Dowty, David. 1979. *Word meaning and Montague Grammar*. Dordrecht: Reidel.

Dryer, Matthew S. 1986. Primary objects, secondary objects and antidative. *Language* 62:808–45.

Dryer, Matthew S. 1997. On the six-way word order typology. *Studies in Language* 21:69–103.

Dryer, Matthew S. 2006. Descriptive theories, explanatory theories, and Basic Linguistic Theory. In *Catching Language: The Standing Challenge of Grammar Writing*, ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 207–234. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Dryer, Matthew S. 2007a. Word order. In *Language Typology and Syntactic Description. Vol. 1: Clause structure*, ed. by Timothy Shopen, 61–131. Cambridge: Cambridge University Press. [2nd edition]

Dryer, Matthew S. 2007b. Clause types. In *Language Typology and Syntactic Description. Vol. 1: Clause Structure*, ed. by Timothy Shopen, 224–275. Cambridge: Cambridge University Press. [2nd edition]

Dryer, Matthew S. 2007c. Noun phrase structure. In *Language Typology and Syntactic Description. Vol. 2: Complex Constructions*, ed. by Timothy Shopen, 151–205. Cambridge: Cambridge University Press. [2nd edition]

Elson, Benjamin, and Velma Pickett. 1988. *Beginning Morphology and Syntax*. Dallas, TX: Summer Institute of Linguistics.

Enfield, Nicholas J. 2006. Heterosemy and the grammar-lexicon trade-off. In *Catching language: The Standing Challenge of Grammar Writing*, ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 297–320. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Evans, Nicholas. 2008. How to write a grammar of an undescribed language: introductory issues. Online: http://www.docstoc.com/docs/2573663/How-to-write-a-grammar-of-an-undescribed-language-introductory-issues.

Evans, Nicholas. 2010. *Dying Words. Endangered Languages and What They Have to Tell Us*. Malden, MA: Wiley-Blackwell.

Evans, Nicholas, and Alan Dench. 2006. Introduction: Catching language. In *Catching Language: The Standing Challenge of Grammar Writing*, ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 1–39. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Evans, Nicholas, and Stephen Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32(5):429–492.

Evans, Nicholas, and Hans-Jürgen Sasse, eds. 2002. *Problems of Polysynthesis*. Berlin: Akademie Verlag.

Everett, Daniel L. 2005. Cultural constraints on grammar and cognition in Pirahã. Another look at the design features of human language. *Current Anthropology* 46(4):621–646.

Everett, Daniel L. 2009a. Pirahã culture and grammar: A response to some criticisms. *Language* 85(2):405–442.

Everett, Daniel L. 2009b. Letter to *Language*. Response to Nevins et al. *Language* 85(4):753.

Farrell, Patrick. 2005. *Grammatical Relations*. (Oxford Surveys in Syntax and Morphology.) Oxford: Oxford University Press.

Fillmore, Charles. 1968. The case for case. In *Universals in Linguistic Theory*, ed. by Emmon Bach and Robert T. Harms, 1–88. New York: Holt, Rinehart & Winston.

Fillmore, Charles, Paul Kay, Laura A. Michaelis, and Ivan A. Sag. 2003. *Construction Grammar*. Stanford: CSLI Publications.

Foley, William A., ed. 1993. *The Role of Theory in Language Description*. Berlin: Mouton de Gruyter.

Foley, William A. 2007. A typology of information packaging in the clause. In *Language Typology and Syntactic Description. Vol. 1: Clause Structure*, ed. by Timothy Shopen, 362–446. Cambridge: Cambridge University Press. [2nd edition]

Foley, William A., and Robert D. Van Valin. 1984. *Functional Syntax and Universal Gramma*r. Cambridge: Cambridge University Press.

Foley, William A., and Robert D. Van Valin. 1985. Information packaging in the clause. In *Language Typology and Syntactic Description. Vol. 1: Clause Structure*, ed. by Timothy Shopen, 282–354. Cambridge: Cambridge University Press.

Fortescue, Michael. 1994. Morphology, polysynthetic. In *The Encyclopedia of Language and Linguistics*, ed. by R. D. Asher, 2600–2602. Oxford: Pergamon.

Foster, Michael K. 1986. Updating the terminology of tense, mood, and aspect in Northern Iroquoian descriptions. *International Journal of American Linguistics* 52(1):65–72.

Fox, Barbara A. 1987. The noun phrase accessibility hierarchy revisited. *Language* 63:856–70.

Gaby, Alice. 2008. Pragmatically case-marked: Non-syntactic functions of the Kuuk Thaayorre ergative suffix. In *Discourse and Grammar in Australian Languages*, ed. by Ilana Mushin and Brett Barker, 111–134. Amsterdam and Philadelphia: John Benjamins.

Geeraerts, Dirk, and Hubert Cuyckens, eds. 2007. *The Oxford Handbook of Cognitive Linguistics*. Oxford and New York: Oxford University Press.

de Geest, Wim, and Y. Putseys. 1984. Sentential Complementation. In Proceedings of the International Conference held at UFSAL, Brussels, June 1983. Dordrecht, The Netherlands: Foris.

Genetti, Carol. 1986. The Development of Subordinators from Postpositions in Bodic Languages. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 12:387–400.

Genetti, Carol. 1991. From postposition to subordinator in Newari. In *Approaches to Grammaticalization* Vol. 2, ed. by Elizabeth Closs Traugott and Bernd Heine, 227–255. Amsterdam and Philadelphia: John Benjamins.

Geniušienė, Emma. 1987. *The Typology of Reflexives*. (Empirical Approaches to language Typology 2.) Berlin: Mouton de Gruyter.

Gerdts, Donna B. 1998. Incorporation. In *The Handbook of Morphology*, ed. by Andrew Spencer and Arnold M. Zwicky, 84–100. Malden, MA: Blackwell.

Gil, David. 2001. Escaping Eurocentrism: fieldwork as a process of unlearning. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 102–132. Cambridge: Cambridge University Press.

Gil, David. 2009. How much grammar does it take to sail a boat? In *Language Complexity as an Evolving Variable* (Studies in the Evolution of Language), ed. by Geoffrey Sampson, David Gil, and Peter Trudgill, 19–33. Oxford: Oxford University Press.

Givón, Talmy. 1978. Definiteness and referentiality. In *Universals of Human Language, Vol. 4*, ed. by Joseph H. Greenberg, Charles A. Ferguson, and Edith A. Moravcsik, 291–330. Stanford, CA: Stanford University Press.

Givón, Talmy. 1980. The binding hierarchy and the typology of complements. *Studies in Language* 4:333–77.

Givón, Talmy, ed. 1983. *Topic Continuity in Discourse: a Quantitative Cross-Language Study*. (Typological Studies in Language 3.) Amsterdam and Philadelphia: John Benjamins.

Goddard, Cliff, ed. 2006. *Ethnopragmatics. Understanding Discourse in Cultural Context*. Berlin and New York: Mouton de Gruyter.

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago, IL: University of Chicago Press.

Good, Jeff. 2006. The Descriptive Grammar as a (Meta)Database. Online: http://emeld.org/workshop/2004/jcgood-paper.html

Green, Georgia M., and Jerry L. Morgan. 2001. *Practical Guide to Syntactic Analysis*, 2nd edition. Chicago, IL: University of Chicago Press.

Greenberg, Joseph H., ed., 1963. *Universals of Language*. Cambridge, MA: MIT Press.

Grimes, Joseph E. 1967. Positional analysis. *Language* 43(2):437–444.

Gruber, Jeffrey. 1965. Studies in Lexical Relations. Unpublished Ph.D. dissertation, MIT

Haas, Mary R. 1969. ‘Exclusive’ and ‘inclusive’: A look at early usage. *International Journal of American Linguistics* 35(1):1–6.

Haegeman, Liliane. 1992. *Theory and Description in Generative Grammar. A Case Study in West Flemish*. Cambridge: Cambridge University Press.

Hagège, Claude. 1982. *La Structure des Langues*. Paris: Presses Universitaires de France.

Haiman, John. 1979. Hua. A Papuan language of New Guinea. In *Languages and Their Status*, ed. by Timothy Shopen, 35–89. Cambridge, MA: Winthrop.

Haiman, John, ed. 1985. *Iconicity in Syntax*. (Typological Studies in Language 6.) Amsterdam and Philadelphia: John Benjamins.

Haiman, John, and Pamela Munro, eds. 1983. *Switch-Reference and Universal Grammar*. (Typological Studies in Language 26.) Amsterdam and Philadelphia: John Benjamins.

Haiman, John, and Sandra A. Thompson, eds. 1988. *Clause Combining in Grammar and Discourse*. (Typological Studies in Language 2.) Amsterdam and Philadelphia: John Benjamins.

Hale, Kenneth L. 1976. The adjoined relative clause in Australian languages. In *Grammatical Categories in Australian Languages*, ed. by R. M. W. Dixon, 78–105 Canberra: Australian Institute of Aboriginal Studies.

Hale, Kenneth L. 1973. A note on subject-object inversion in Navajo. In *Issues in Linguistics. Papers in Honor of Henry and Renee Kahane*, ed. by Braj Kachru et al., 300–309. Urbana, IL: University of Illinois Press.

Hale, Kenneth L. 1983. Warlpiri and the grammar of non-configurational languages. *Natural Language and Linguistic Theory* 1(1):5–47.

Halliday, M. A. K. 1994. *An Introduction to Functional Grammar*, 2nd ed. London: Edward Arnold.

Hanks, William F. 2005. Explorations in the deictic field. *Current Anthropology* 46(2):191–220.

Harris, Alice C. 2002. *Endoclitics and the Origins of Udi Morphosyntax*. Oxford: Oxford University Press.

Harris, Zellig S. 1952. Discourse analysis. *Language* 28(1):1–30.

Harris, Zellig S. 1957. Cooccurrence and transformation in linguistic structure. *Language* 33(3):283–340.

Haspelmath, Martin. 1995. The converb as a cross-linguistically valid category. In *Converbs in Cross-Linguistic Perspective*, ed. by Martin Haspelmath and Ekkehard König, 1–55. Berlin and New York: Mouton de Gruyter.

Haspelmath, Martin. 1996. Word-class-changing inflection and morphological theory. In *Yearbook of Morphology*, ed. by Geert Booij and Jaap van Marle, 43–44. Dordrecht, The Netherlands: Kluwer.

Haspelmath, Martin. 1997. *Indefinite Pronouns*. (Oxford Studies in Typology and Linguistic Theory) Oxford: Oxford University Press.

Haspelmath, Martin. 2007. Coordination. In *Language Typology and Syntactic Description. Vol. 2: Complex Constructions*, ed. by Timothy Shopen, 1–51. Cambridge: Cambridge University Press. [2nd edition]

Haspelmath, Martin, and Ekkehard König, eds. 1995. *Converbs in Cross-Linguistic Perspective. Structure and meaning of Adverbial Verb Forms – Adverbial Participles, Gerunds*. Berlin and New York: Mouton de Gruyter.

Haspelmath, Martin, and Andrea Simms. 2009. *Understanding Morphology*. (2nd edition.) Oxford: Oxford University Press.

Haviland, John B. 1979. How to talk to your brother-in-law in Guugu Yimidhirr. In *Languages and Their Speakers*, ed. by Timothy Shopen, 161–239. Cambridge, MA: Winthrop.

Hawkins, John A. 1983. *Word Order Universals*. New York: Academic.

Hawkins, John A. 1985. Complementary methods in universal grammar: a reply to Coopmans. *Language* 61(3):569–87.

Hawkins, John A. 1989. Competence and performance in the explanation of language universals. In *Essays on Grammatical Theory and Universal Grammar*, ed. by Doug Arnold, Martin Atkinson, Jacques Durand, Claire Grover, and Louisa Sadler, 119–152. Oxford: Clarendon Press.

Heine, Bernd. 1993. *Auxiliaries: Cognitive Forces and Grammaticalization*. Oxford: Oxford University Press.

Helmbrecht, Johannes, and Christian Lehmann. 2008. Hočank's challenge to morphological theory. In *Lessons from Documented Endangered Languages*, ed. by K. David Harrison, David S. Rood, and Arienne Dwyer, 271–315. (Typological Studies in Language 78.) Amsterdam and Philadelphia: John Benjamins.

Hengeveld, Kees. 1993. *Non-Verbal Predication: Theory, Typology, Diachrony*. Berlin: Mouton de Gruyter.

Himmelmann, Nikolaus P. 2006a. How to miss a paradigm or two: Multifunctional *ma-* in Tagalog. In *Catching Language: The Standing Challenge of Grammar Writing*, ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 487–526. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Himmelmann, Nikolaus P. 2006b. The challenges of segmenting spoken language. In *Essentials of Language Documentation*, ed. by Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel, 253–274. (Trends in Linguistics. Studies and Monographs 178.) Berlin and New York: Mouton de Gruyter.

Hinton, Leanne, Johanna Nichols, and John Ohala, eds. 1994. *Sound Symbolism*. Cambridge: Cambridge University Press.

Hockett, Charles F. 1954. Two models of grammatical description. *Word* 10:210–34.

Hoeksema, Jack. 1985. *Categorial Morphology*. (Outstanding Dissertations in Linguistics Series.) New York: Garland Press.

de Hoop, Helen, and Peter de Swart. 2008. *Differential Subject Marking*. Dordrecht, The Netherlands: Springer.

Hopper, Paul J., and Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56:251–99.

Hopper, Paul J., and Elizabeth Closs Traugott. 1993. *Grammaticalization*. (Cambridge Textbooks in Linguistics.) Cambridge: Cambridge University Press.

Huang, Yan. 2000. *Anaphora*: *A Cross-Linguistic Study*. (Oxford Studies in Typology and Linguistic Theory). Oxford: Oxford University Press.

Jackendoff, Ray. 1977. *X' Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.

Jacobsen, William H., Jr. 1967. Switch-reference in Hokan-Coahuiltecan. In *Studies in Southwestern Ethnolinguistics: Meaning and History in the Languages of the American Southwest*, ed. by Dell Hymes, 283–63. The Hague: Mouton.

Jacobsen, William H., Jr. 1979. Noun and verb in Nootkan. In *The Victoria Conference on Northwestern Languages*, ed. by Barbara S. Efrat, 83–155. (British Columbia Provincial Museum Heritage Record No. 4.) Victoria, B.C.: British Columbia Provincial Museum.

Janssen, Theo A. J. M., and Wim van der Wurff, eds. 1996. *Reported Speech. Forms and Functions of the Verb.* (Pragmatics and beyond New Series 43.) Amsterdam and Philadelphia: John Benjamins.

Jendraschek, Gerd. 2001. Semantic and structural properties of Turkish ideophones. *Turkic Languages* 5:88–103.

Jelinek, Eloise. 1984. Empty categories, case and configurationality. *Natural Language and Linguistic Theory* 2:39–76.

Johnson, Rodney C. 1992. *The limits of grammar: Syntax and lexicon in spoken Burmese*. Unpublished Ph.D. Dissertation, University of Michigan.

Kärkäinen, Elise, Marja-Leena Sorjonen, and Marja Liisa Helasvuo. 2007. Discourse structure. In *Language Typology and Syntactic Description. Vol. 2: Complex Constructions*, ed. by Timothy Shopen, 301–371. Cambridge: Cambridge University Press. [2nd edition]

Kazenin, Konstantin I. 1994. Split syntactic ergativity: toward an implicational hierarchy. *Sprachtypologie und Universalienforschung* 47:78–98.

Keenan, Edward L. 1985a. Passive in the world's languages. In *Language Typology and Syntactic Description. Vol. 1: Clause Structure*, ed. by Timothy Shopen, 243–281. Cambridge: Cambridge University Press.

Keenan, Edward L. 1985b. Relative clauses. In *Language Typology and Syntactic Description. Vol. 2: Complex Constructions*, ed. by Timothy Shopen, 141–170. Cambridge: Cambridge University Press.

Keenan, Edward L., and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8(1):63–99.

Keenan, Edward L., and Matthew S. Dryer. 2007. Passive in the world's languages. In *Language Typology and Syntactic Description. Vol. 1: Clause Structure*, ed. by Timothy Shopen, 325–361. Cambridge: Cambridge University Press. [2nd edition]

Keenan, Edward L., and Elinor Ochs. 1979. Becoming a competent speaker of Malagasy. In *Languages and Their Speakers*, ed. by Timothy Shopen, 113–58. Cambridge, MA: Winthrop.

Kemmer, Suzanne. 1993. *The Middle Voice*. (Typological Studies in Language 23.) Amsterdam and Philadelphia: John Benjamins.

Kinkade, M. Dale. 1983. Salish evidence against the universality of "noun" and "verb". *Lingua* 60:25–40.

Kiss, Katalin É, ed. 1995. *Discourse Configurational Languages*. (Oxford Studies in Comparative Syntax.) New York and Oxford: Oxford University Press.

Kiss, Katalin É. 1998. Identificational focus versus information focus. *Language* 74(2):245–73.

Klaiman, Mimi H. 1991. *Grammatical Voice*. Cambridge: Cambridge University Press.

Klaiman, Mimi H. 1992. *Inverse languages*. Lingua 88(3/4):227–61.

Klavans, Judith L. 1985. The independence of syntax and phonology in cliticization. *Language* 61:95–120.

König, Ekkehard. 1995. The meaning of converb constructions. In *Converbs in Cross-Linguistic Perspective*, ed. by Martin Haspelmath and Ekkehard König, 57–95. Berlin and New York: Mouton de Gruyter.

König, Ekkehard, and Volker Gast, eds. 2008. *Reciprocals and Reflexives. Theoretical and Typological Explorations*. Berlin and New York: Mouton de Gruyter.

König, Ekkehard, and Peter Siemund. 2007. Speech act distinctions in grammar. In *Language Typology and Syntactic Description. Vol. 1: Clause Structure*, ed. by Timothy Shopen, 276–342. Cambridge: Cambridge University Press. [2nd edition]

Koptjevskaja-Tamm, Maria. 1993. *Nominalizations*. London: Routledge.

Kortmann, B. 1997. *Adverbial Subordination. A Typology and History of Adverbial Subordinators Based on European Languages*. Berlin: Mouton de Gruyter.

Kroeger, Paul. 2004. *Analyzing Syntax. A Lexical-Functional Approach*. Cambridge: Cambridge University Press.

Kuipers, Aert. 1968. The categories verb-noun and transitive-intransitive in English and Squamish. *Lingua* 21:610–26.

Langacker, Ronald W. 1991. Cognitive grammar. In *Linguistic Theory and Grammatical Description. Nine Current Approaches*, (Current Issues in Linguistic Theory 75) ed. by Flip G. Droste and John E. Joseph, 275–306. Amsterdam and Philadelphia: John Benjamins.

LaPolla, Randy J. 1992. Anti-ergative marking in Tibeto-Burman. *Linguistics of the Tibeto-Burman Area* 15(1):1–9.

LaPolla, Randy J., and Dory Poa. 2006. On describing word order. In *Catching Language: The Standing Challenge of Grammar Writing*, ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 269–295. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Lehmann, Christian. 1984. *Der Relativsatz: Typologie seiner Strukturen, Theorie seiner Funktionen, Kompendium seiner Grammatik*. Tübingen: Gunter Narr.

Levin, Beth C. 1993. *English Verb Classes and Alternations*. Chicago, IL: University of Chicago Press.

Levin, Erik. 2007. A reanalysis of the Aymara verb. In *Language Endangerment and Endangered Languages. Linguistic and Anthropological Studies with Special Emphasis on the Languages and Cultures of the Andean-Amazonian Border Area*. (Indigenous Languages of Latin America 5.), ed. by W. Leo Wetzels, 428–440. Leiden: CNWS.

Levinson, Stephen C. 1983. *Pragmatics*. (Cambridge Textbooks in Linguistics.) Cambridge: Cambridge University Press.

Levinson, Stephen C. 2003. *Space in Language and Cognition: Explorations in Cognitive Diversity*. (Language, Culture and Cognition 5.) Cambridge: Cambridge University Press.

Levinson, Stephen C., and David Wilkins, eds. 2006. *Grammars of Space: Explorations in Cognitive Diversity*. Cambridge: Cambridge University Press.

Li, Charles N., ed. 1976. *Subject and Topic*. New York: Academic.

Li, Charles N., and Sandra A. Thompson. 1981. *Mandarin Chinese. A Functional Reference Grammar*. Berkeley, CA and Los Angeles, CA: University of California Press.

Lindsey, G., and Janine Scancarelli. 1985. Where have all the adjectives come from? The case of Cherokee. *Proceedings of the Annual Meeting of the Berkeley Linguistic Society* 11:207–215.

Lockwood, David G. 2002. *Syntactic Analysis and Description. A Constructional Approach*. London and New York: Continuum.

Longacre, Robert E. 1985. Sentences as combinations of clauses. In *Language Typology and Syntactic Description. Vol. 2: Complex Constructions*, ed. by Timothy Shopen, 235–286. Cambridge: Cambridge University Press.

Longacre, Robert E. 2007. Sentences as combinations of clauses. In *Language Typology and Syntactic Description. Vol. 2: Complex Constructions*, ed. by Timothy Shopen, 372–420. Cambridge: Cambridge University Press. [2nd edition]

Lord, Carol. 1993. *Historical Change in Serial Verb Constructions*. (Typological Studies in Language 26.) Amsterdam and Philadelphia: John Benjamins.

Luraghi, Silvia, and Claudia Parodi. 2008. *Key Terms in Syntax and Syntactic Theory*. London and New York: Continuum.

Lynch, John. 1998. *Pacific languages: An Introduction*. Honolulu: University of Hawai'i Press.

Lyons, Christopher. 1999. *Definiteness*. (Cambridge Textbooks in Linguistics.) Cambridge: Cambridge University Press.

Lyovin, Anatole V. 1997. *An Introduction to the Languages of the World*. New York and Oxford: Oxford University Press.

Malchukov, A., and A. Spencer. 2008. Typology of case systems, parameters of variation. In *Handbook of Case*, ed. by A. Malchukov and A. Spencer, 651–667. Oxford: Oxford University Press.

Mallinson, Graham, and Barry J. Blake. 1981. *Language Typology: Cross-Cultural Studies in Syntax*. Amsterdam: North-Holland.

Marchand, Hans. 1969. *The Categories and Types of Present-Day English Word-Formation. A Synchronic-Diachronic Approach.* 2nd edition. Munich: Beck'sche Verlagsbuchhandlung.

Martinet, André. 1964. *Elements of General Linguistics*. With a Foreword by L. R. Palmer. Chicago, IL: University of Chicago Press.

Mattissen, Johanna. 2004. A structural typology of polysynthesis. *Word* 55:189–216.

McCarthy, John J. 1981. A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* 12:373–418.

McCawley, James D. 1995. English. In *Syntax. An International Handbook of Contemporary Research*, ed. by Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld and Theo Vennemann. Vol. 2., 1319–1347. Berlin and New York: Walter de Gruyter.

McCawley, James D. 1998. *The Syntactic Phenomena of English*. 2nd ed. Chicago, IL: Chicago University Press.

Mel'chuk, Igor. 2006. Calculus of possibilities as a technique in linguistic typology. In *Catching Language: The Standing Challenge of Grammar Writing*, ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 171–205. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Mel'chuk, Igor, and David Beck. 2006. *Aspects of the Theory of Morphology*. Berlin and New York: Mouton de Gruyter.

Merrifield, William R., Constance M. Naish, Calvin R. Rensch, and Gillian Story, eds. 2003. *Laboratory Manual for Morphology and Syntax*. 7th edition. Dallas, TX: SIL International.

Mithun, Marianne. 1984. The evolution of noun incorporation. *Language* 60:847–95.

Mithun, Marianne. 1991. Active/agentive case marking and its motivations. *Language* 67:510–46.

Mithun, Marianne. 1992. Is basic word order universal? In *Pragmatics of Word order Flexibility*, ed. by Doris Payne, 15–62. (Typological Studies in language 22.) Amsterdam and Philadelphia: John Benjamins.

Mosel, Ulrike. 2006. Grammaticography: The art and craft of writing grammars. In *Catching Language: The Standing Challenge of Grammar Writing*, ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 11–68. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Mulder, Jan. 1989. The viability of the notion of subject in Coast Tsimshian. *Canadian Journal of Linguistics* 34(2):129–44.

Munro, Pamela. 2007. From parts of speech to the grammar. In *Perspectives on Grammar Writing*, ed. by Thomas E. Payne and David J. Weber, 72–111. (Benjamins Current Topics 11.) Amsterdam and Philadelphia: John Benjamins.

Mushin, Ilana, and Brett Baker, eds. 2008. *Discourse and Grammar in Australian Languages*. (Studies in Language Companion Series 104.) Amsterdam and Philadelphia: John Benjamins.

Muysken, Pieter. 1986. Approaches to affix order. *Linguistics* 24(3):629–43.

Nedjalkov, Igor' V. 1998. Converbs in the languages of Europe. In *Adverbial Constructions in the Languages of Europe*, ed. by Johan van der Auwera and Donall P. O'Baoill, 421–455. Berlin: Mouton de Gruyter.

Nevins, Andrew, David Pesetsky, and Cilene Rodrigues. 2009a. Pirahã exceptionality: A reassessment. *Language* 85(2):355–404.

Nevins, Andrew, David Pesetsky, and Cilene Rodrigues. 2009b. Evidence and argumentation: A Reply to Everett (2009). *Language* 85(3):671–681.

Newmeyer, Frederick J. 2001. Deconstructing grammaticalization. *Language Sciences* 23.187–229.

Nichols, Johanna. 1986. Head-marking and dependent-marking grammar. *Language* 62(1):56–119.

Nichols, Johanna, and Anthony C. Woodbury, eds. 1985. *Grammar Inside and Outside the Clause: Some Approaches to Theory from the Field*. Cambridge: Cambridge University Press.

Nida, Eugene A. 1949. *Morphology. The Descriptive Analysis of Words*. Ann Arbor, MI: The University of Michigan Press.

Noonan, Michael. 1985. Complementation. In *Language Typology and Syntactic Description. Vol. 2: Complex Constructions*, ed. by Timothy Shopen, 42–140. Cambridge: Cambridge University Press.

Noonan, Michael. 2007. Complementation. In *Language Typology and Syntactic Description. Vol. 2: Complex Constructions*, ed. by Timothy Shopen, 52–150. Cambridge: Cambridge University Press. [2nd edition]

O'Connor, Loretta. 2007. *Motion, Transfer and Transformation. The Grammar of Change in Lowland Chontal*. Amsterdam and Philadelphia: John Benjamins.

Olson, Michael L. 1981. *Barai clause junctures: toward a functional theory of interclausal relations*. Unpublished Ph.D. Dissertation, Australian National University, Canberra.

Palmer, F. R. 1986. *Mood and Modality*. (Cambridge Textbooks in Linguistics.) Cambridge: Cambridge University Press.

Palmer, F. R. 1994. *Grammatical Roles and Relations*. (Cambridge Textbooks in Linguistics.) Cambridge: Cambridge University Press.

Pawley, Andrew K. 1987. Encoding events in Kalam and English: Different logics for reporting experience. In *Coherence and Grounding in Discourse*. (Typological Studies in Language 11.), ed. by Russell S. Tomlin, 329–360. Amsterdam and Philadelphia: John Benjamins.

Pawley, Andrew K. 1993. A language which defies description by ordinary means. In *The Role of Theory in Language Description*, ed. by William A. Foley, 87–129. Berlin: Mouton de Gruyter.

Payne, Doris L. 1990. *The pragmatics of word order: Typological dimensions of verb-initial languages*. Berlin: Walter de Gruyter.

Payne, Doris L., ed. 1992. *Pragmatics of Word order Flexibility*. (Typological Studies in Language 22.) Amsterdam and Philadelphia: John Benjamins.

Payne, John R. 1985a. Negation. In *Language Typology and Syntactic Description. Vol. 1: Clause Structure*, ed. by Timothy Shopen, 197–242. Cambridge: Cambridge University Press.

Payne, John R. 1985b. Complex phrases and complex sentences. In *Language Typology and Syntactic Description. Vol. 2: Complex Constructions*, ed. by Timothy Shopen, 3–41. Cambridge: Cambridge University Press.

Payne, Thomas E. 1991. Categorized and prioritized bibliography on language typology. *Notes on Linguistics* 54(43):7–9.

Payne, Thomas E. 1997. *Describing Morphosyntax: a Guide for Field Linguists*. Cambridge: Cambridge University Press.

Peck, Charles W. 1981. *A Survey of Grammatical Structures*. 2nd ed. Dallas, TX: SIL International.

Perlmutter, David M. 1978. Impersonal passives and the unaccusative hypothesis. *Proceedings of the Berkeley Linguistics Society* 4:157–89.

Peterson, David A. 2007. *Applicative Constructions*. (Oxford Studies in Typology and Linguistic Theory) Oxford: Oxford University Press.

Pike, Kenneth L. 1947. Grammatical prerequisites to phonemic analysis. *Word* 3:55–72.

Pike, Kenneth L. 1952. More on grammatical prerequisites. *Word* 8:106–121.

Plank, Frans, ed. 1995. *Double Case: Agreement by Suffixaufname*. New York: Oxford University Press.

Postal, Paul M. 1988. Topic…comment. Advances in linguistic rhetoric. *Natural Language and Linguistic Theory* 6:129–37.

Prince, Ellen F. 1981. Toward a taxonomy of given-new information. In *Radical Pragmatics*, ed. by Peter Cole, 223–255. New York: Academic.

Pustet, Regina. 2003. *Copulas. Universals in the Categorization of the Lexicon.* (Oxford Studies in Typology and Linguistic Theory) Oxford: Oxford University Press.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London and New York: Longman.

Radford, Andrew. 1988. *Transformational Grammar: A First Course*. Cambridge: Cambridge University Press.

Ransom, Evelyn. 1986. *Complementation: Its Meaning and Forms*. (Typological Studies in Language 10.) Amsterdam and Philadelphia: John Benjamins.

Reesink, Ger P. 1983. Switch reference and topicality hierarchies. *Studies in Language* 7(2):215–246.

de Reuse, Willem J. 1993. Noun incorporation. In *The Encyclopedia of Language and Linguistics*, ed. by R. E. Asher, 2842–2847. Oxford: Pergamon Press and Aberdeen University Press.

de Reuse, Willem J. 1994. *Siberian Yupik Eskimo. The Language and Its Contacts with Chukchi*. Salt Lake City, UT: University of Utah Press.

de Reuse, Willem J. 2006. Polysynthetic language: Central Siberian Yupik Eskimo. In *Encyclopedia of Language and Linguistics*, 2nd edition, ed. by Keith Brown, 745–748. Oxford: Elsevier.

de Reuse, Willem J. 2009. Polysynthesis as a typological feature: An attempt at a characterization from Eskimo and Athabaskan perspectives. In *Variations on Polysynthesis, Proceedings of the Linguistics Session of the 15th International Inuit Studies Conference* (Typological Studies in Language 86), edited by Marc-Antoine Mahieu and Nicole Tersis, 19–34. Amsterdam and Philadelphia: John Benjamins.

Rice, Keren. 1989. *A Grammar of Slave*. (Mouton Grammar Library 5.) Berlin and New York: Mouton de Gruyter.

Rice, Keren. 2006. Let the language tell its story? The role of linguistic theory in writing grammars. In *Catching language: The Standing Challenge of Grammar Writing*, ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 235–268. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Rice, Keren. 2007. A Typology of good grammars. In *Perspectives on Grammar Writing*, ed. by Thomas E. Payne and David J. Weber, 143–171. (Benjamins Current Topics 11.) Amsterdam and Philadelphia: John Benjamins.

Rijkhoff, Jan. 2002. *The Noun Phrase*. (Oxford Studies in Typology and Linguistic Theory). Oxford: Oxford University Press.

Romero-Figueroa, Andrés. 1985. OSV as the basic order in Warao. *Linguistics* 23(1):105–21.

Rosch, Eleanor. 1975. Universals and cultural specifics in human categorization. In *Cross-Cultural Perspectives on Learning*, ed. by Richard W. Brislin, Stephen Bochner, and Walter J. Lonner, 177–206. New York: Sage Publications.

Sadock, Jerrold M. 1991. *Autolexical Syntax. A Theory of Parallel Grammatical Representations*. Chicago, IL and London: University of Chicago Press.

Sadock, Jerrold M. 1996. PIFL: The Principle of Information Free Linguistics. *Chicago Linguistic Society (Parasession on Theory and Data in Linguistics)* 32:133–138.

Sadock, Jerrold M., and Arnold M. Zwicky 1985. Speech act distinctions in syntax. In *Language Typology and Syntactic Description. Vol. 1: Clause Structure*, ed. by Timothy Shopen, 155–196. Cambridge: Cambridge University Press.

Sampson, Geoffrey. 2009. A linguistic axiom challenged. In *Language Complexity as an Evolving Variable* (Studies in the Evolution of Language), ed. by Geoffrey Sampson, David Gil, and Peter Trudgill, 1–18. Oxford: Oxford University Press.

Sapir, Edward. 1917. Review of C. C. Uhlenbeck, Het passieve karakter van het verbum transitivum of van het verbum actionis in talen van Noord-Amerika. *International Journal of American Linguistics* 1:82–86.

Sapir, Edward. 1921. *Language: An Introduction to the Study of Speech*. New York: Harcourt Brace.

Sasse, Hans-Jürgen. 2002. Recent activity in the theory of aspect: Accomplishments, achievements, or just non-progressive state? *Linguistic Typology* 6(2):199–271.

Schachter, Paul. 1985. Parts-of-speech systems. In *Language Typology and Syntactic Description. Vol. 1: Clause Structure*, ed. by Timothy Shopen, 3–61. Cambridge: Cambridge University Press.

Schachter, Paul, and Timothy Shopen. 2007. Parts-of-speech systems. In *Language Typology and Syntactic Description. Vol. 1: Clause Structure*, ed. by Timothy Shopen, 1–60. Cambridge: Cambridge University Press. [2nd edition]

Schiffrin, Deborah. 1994. *Approaches to Discourse*. (Blackwell Textbooks in Linguistics.) Oxford and Cambridge: Blackwell.

Schultze-Berndt, Eva. 2006. Taking a closer look at function verbs: Lexicon, grammar, or both? In *Catching Language: The Standing Challenge of Grammar Writing*, ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 359–391. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Schulze, Wolfgang. 2005. Grammars for East Caucasian. *Anthropological Linguistics* 47(3):321–352.

Seibert, Uwe. 2000. Writing a reference grammar. Dept. of Languages and Linguistics, University of Jos, Nigeria. Online: http://www.uiowa.edu/intlinet/unijos/nigonnet/nlp/refgram1.htm

Selkirk, O. Elizabeth. 1982. *The Syntax of Words*. (Linguistic Inquiry Monograph 7.) Cambridge, MA: MIT Press.

Sen-Gupta, Sunil. 1978. Common cognitive elements in combined verbs (Bengali, Khmer, Bahasa Indonesian [sic], and Vietnamese). In *Approaches to Languages: Anthropological Issues*, ed. by William C. McCormack and Stephen A. Wurm, 363–367. The Hague: Mouton.

Shibatani, Masayoshi, ed. 1988. *Passive and Voice*. (Typological Studies in Language 16.) Amsterdam and Philadelphia: John Benjamins.

Shopen, Timothy, ed. 1985. *Language Typology and Syntactic Description. vol. 1: Clause Structure; vol. 2: Complex Constructions; vol. 3: Grammatical Categories and the Lexicon*. Cambridge: Cambridge University Press.

Shopen, Timothy, ed. 2007. *Language Typology and Syntactic Description. vol. 1: Clause Structure; vol. 2: Complex Constructions; vol. 3: Grammatical Categories and the Lexicon*. Cambridge: Cambridge University Press. [2nd edition]

Siewierska, Anna. 1984. *The Passive: A Cross-Linguistic Analysis*. London: Croom Helm.

Siewierska, Anna. 2004. *Person*. (Cambridge Textbooks in Linguistics.) Cambridge: Cambridge University Press.

Silver, Shirley, and Wick R, Miller. 1997. *American Indian Languages. Social and Cultural Contexts*. Tucson, AZ: University of Arizona Press.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In *Grammatical Categories in Australian Languages*, ed. by R. M. W. Dixon, 112–71. Canberra: Australian Institute of Aboriginal Studies.

Simpson, Jane, and Meg Withgott. 1986. Pronominal Clitic Clusters and Templates. In *Syntax and semantics 19: The Syntax of Pronominal Clitics*, ed. by Hagit Borer, 149–74. New York: Academic.

Song, Jae Jung. 2001. *Linguistic Typology: Morphology and Syntax*. London: Longman.

Spencer, Andrew, and Arnold M. Zwicky, eds. 1998. *The Handbook of Morphology*. Malden, MA: Blackwell.

Stassen, Leon. 1985. *Comparison and Universal Grammar*. Oxford: Basil Blackwell.

Stassen, Leon. 1997. *Intransitive Predication*. (Oxford Studies in Typology and Linguistic Theory) Oxford: Oxford University Press

Talmy, Leonard. 1985. Lexicalization patterns: semantic structure in lexical forms. In *Language Typology and Syntactic Description. Vol. 3: Grammatical Categories and the Lexicon*, ed. by Timothy Shopen, 57–149. Cambridge: Cambridge University Press.

Talmy, Leonard. 2000. *Toward a Cognitive Semantics*. *Vol I: Concept Structuring Systems, Vol II: Typology and Process in Concept Structuring*. Cambridge, MA: MIT Press.

Talmy, Leonard. 2007. Lexical Typologies. In *Language Typology and Syntactic Description. Vol. 3: Grammatical Categories and the Lexicon*, ed. by Timothy Shopen, 66–168. Cambridge: Cambridge University Press. [2nd edition].

Thompson, Lawrence C. 1979. The control system: A major category in the grammar of Salishan languages. In *The Victoria Conference on Northwestern Languages*, ed. by Barbara S. Efrat, 156–176. (British Columbia Provincial Museum Heritage Record No. 4.) Victoria, B.C.: British Columbia Provincial Museum.

Thompson, Lawrence C. 1985. Control in Salish grammar. In *Relational Typology*, ed. by Frans Plank, 391–428. (Trends in Linguistics. Studies and Monographs 28.) The Hague: Mouton.

Thompson, Sandra A., and Robert E. Longacre. 1985. Adverbial clauses. In *Language Typology and Syntactic Description. Vol. 2: Complex Constructions*, ed. by Timothy Shopen, 171–234. Cambridge: Cambridge University Press.

Thompson, Sandra A., Robert E. Longacre, and Shin Ja J. Hwang. 2007. Adverbial clauses. In *Language Typology and Syntactic Description. Vol. 2: Complex Constructions*, ed. by Timothy Shopen, 237–300. Cambridge: Cambridge University Press. [2nd edition]

Timberlake, Allan. 2007. Aspect, tense, mood. In *Language Typology and Syntactic Description. Vol. 3: Grammatical Categories and the Lexicon*, ed. by Timothy Shopen, 280–333. Cambridge: Cambridge University Press. [2nd edition]

Turner, Blaine. 1986. *A Teaching Grammar of the Manam Language*. (Data Papers on Papua New Guinea Languages, Vol. 34.) Ukarumpa, Papua New Guinea: Summer Institute of Linguistics.

Ultan, Russell. 1978a. Some general characteristics of interrogative systems. In *Universals of Human Language, Vol. IV: Syntax*, ed. by Joseph H. Greenberg, 211–46. Stanford, CA: Stanford University Press.

Ultan, Russell. 1978b. Toward a typology of substantival possession. In *Universals of Human Language, Vol. IV: Syntax*, ed. by Joseph H. Greenberg, 11–49. Stanford, CA: Stanford University Press.

Uray, Géza. 1954. Duplication, gemination, and triplication in Tibetan. *Acta Orientalia Academiae Scientiarium Hungaricae* 4(1/3):177–244.

Van Valin, Robert D. Jr., and Randy LaPolla. 1997. *Syntax: Structure, Meaning and Function*. Cambridge: Cambridge University Press

Vendler, Zeno. 1967. *Philosophy in Linguistics*. Ithaca: Cornell University Press.

Vogel, Petra M., and Bernard Comrie, eds. 2000. *Approaches to the Typology of Word Classes*. Berlin and New York: Mouton de Gruyter.

Welmers, William E. 1973. *African Language Structures*. Berkeley, CA and Los Angeles, CA: University of California Press.

Wetzer, Harrie. 1996. The *Typology of Adjectival Predication*. Berlin and New York: Mouton de Gruyter.

Whaley, Lindsay. 1997. *Introduction to Typology: The Unity and Diversity of Language*. Thousand Oaks, CA: Sage Publications.

Wierzbicka, Anna. 1976. Particles and linguistic relativity. *International Review of Slavic Linguistics* 1(1/2):327–67.

Wierzbicka, Anna. 1986a. A semantic metalanguage for the description and comparison of illocutionary meanings. *Journal of Pragmatics* 10:67–107.

Wierzbicka, Anna, ed. 1986b. Special issue on 'particles'. *Journal of Pragmatics* 10(5):519–643.

Wierzbicka, Anna. 1988. *The Semantics of Grammar*. Amsterdam and Philadelphia: John Benjamins.

Wierzbicka, Anna. 1994. Semantics and epistemology: the meaning of 'evidentials' in a cross-linguistic perspective. *Language Sciences* 16(1):81–137.

Wilhelm, Andrea. 2007. *Telicity and Durativity. A Study of Aspect in Dëne Sųłiné (Chipewyan) and German*. New York and London: Routledge.

Woodbury, Anthony C. 1993. A defense of the proposition: "When a language dies, a culture dies." (Proceedings of the first annual symposium about language and society-Austin (SALSA).) *Texas Linguistic Forum* 33:101–129.

Xrakovskij, Victor S., ed. 2001. *Typology of Imperative Constructions*. (Lincom Studies in Theoretical Linguistics 9.) Munich: Lincom-Europa.

Zúñiga, Fernando. 2007. *Mapudungun. El habla mapuche. Introducción a la lengua mapuche, con notas comparativas y un CD*. Santiago de Chile: Centro de Estudios Públicos.

Zwicky, Arnold M. 1985. Heads. *Journal of Linguistics* 21:1–29.

Zwicky, Arnold M., and Geoffrey K. Pullum. 1983. Cliticization versus inflection: English *n't*. *Language* 59:502–13.

Chapter 12
Grammar Gathering Techniques

I have no data yet. It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts.

Sherlock Holmes in 'A Scandal in Bohemia', by Sir Arthur Conan Doyle.

We must be cautious, we musn't [sic] manufacture facts to fit a theory."

Detective-Inspector Napoleon Bonaparte in 'The Mystery of Swordfish Reef' by Arthur Upfield.

(Both quoted in the Research Centre for Linguistic Typology Newsletter (La Trobe University, Bundoora, Australia, February 2003))

12.1 Introduction

This chapter focuses on grammar gathering techniques. By "grammar" we mean the morphology and syntax of a language. Phonology and phonetics data gathering techniques were addressed in Chapter 10. In this introduction, we present preliminary questions about grammar gathering techniques (Section 12.1.1), and then discuss issues of terminology and classification regarding grammar gathering tasks (Section 12.1.2). Then, as in Chapter 11, we divide the discussion into morphosyntax, morphology, and syntax. Morphosyntax gathering techniques (Section 12.2) are those that can be applied equally well to morphology and to syntax; morphology gathering techniques (Section 12.3) apply to morphology only; and syntax gathering techniques (Section 12.4) apply to syntax only. Section 12.5 concludes with some general thoughts about grammar gathering tasks.

S.L. Chelliah and W.J. de Reuse, *Handbook of Descriptive Linguistic Fieldwork*,
DOI 10.1007/978-90-481-9026-3_12,

12.1.1 Three Questions on Grammar Gathering

Answers to three main questions define a researcher's position on how and why data is gathered for linguistic analysis. To choose the methods of data gathering, these positions must first be understood. These questions are:

1. Should data gathering be theory-driven or data-driven?
2. Is data gathering performed in a predictable linear fashion?
3. What data should be gathered?

12.1.1.1 Theory-Driven or Data-Driven Data Gathering

Under one view, data collection should begin with a particular theoretical orientation. Vaux et al. (2007) describe themselves as generativists, and approach fieldwork from this perspective. They rephrase Morris Halle as follows:

> Data on their own are meaningless; it is the theoretical framework which dictates what facts are interesting and what facts are not. The theory, moreover, tells us what questions to ask; without a well-articulated linguistic theory, the field worker (like most pre-nineteenth century scholars) is limited to disconnected anecdotal observations.
>
> (Vaux et al. 2007:2)

However, a second view is that data gathering should precede theorizing. This view is reflected in the two quotations under the chapter title.

In our opinion, both of these views have merit. Descriptive linguistic fieldwork is not and cannot be based solely on deductive reasoning, nor can it be based solely on inductive generalizations drawn from the observation of facts. Successful descriptive fieldwork involves a constant back-and-forth between data observation on the one hand, and theoretical, deductive modeling on the other – without ever losing track of the primacy of observable data.

As pointed out already in our discussion of theory in Section 11.4.2, no linguist is a theoretical "tabula rasa", nor is it desirable (or even possible) to be one. Bowern (2008:11) also remarks on the Sherlock Holmes quote given at the beginning of this chapter – that no linguist approaches data without any theoretical presuppositions at all. Thus each fieldwork data gathering endeavor has some theoretical underpinning. In addition, theory-driven research can be useful in determining what data to collect, especially with regard to more nuanced grammatical points, as further discussed in Rice (2001:240–245).

Linguistics is unlike mathematics or logic. Mathematics and logic are mainly deductive sciences – sciences where isolated observations are unimportant or even trivial if they do not fit in the theory. As pointed out already in Section 11.4.2, a deductive approach belittles the importance of observational adequacy. Because of the generativists' conception of linguistics as a primarily deductive science, observational adequacy appears to be less important: observation is pre-theoretical, in the sense that there is no "observational theory" parallel to Chomsky's descriptive and explanatory theories. But as we will see again in Section 12.1.1.3 below, observation is crucial in descriptive linguistic fieldwork.[1]

Ideally, then, grammar gathering must be neither only data-driven nor only theory-driven, but must be both simultaneously.

12.1.1.2 Directionality in Data Gathering

It should be obvious that a fieldworker first begins to examine and analyze smaller morphosyntactic structures, and then moves on to examine and analyze larger morphosyntactic structures. However, as Dixon (2007, 2010a) points out, and Krishnamurti (2007:58), Hyman (2001:30–31), and Hale (2001:84) also imply, this does not mean that there is a directionality in fieldwork. That is, it is not the case that one can complete one data gathering technique before moving sequentially on to another. The problem with the "discovery procedures" of the American structuralists (Longacre 1964) was not so much that there were no discovery procedures, but rather that one was supposed to discover morphosyntactic facts in a certain order.[2] That simply does not work. For example, syntax data cannot be gathered before morphology data, or vice versa, any more than noun and verb morphology can be gathered separately or sequentially.

In actual practice, the fieldworker will try to divide large problems into smaller ones, and s/he will at least try to start from "simple" constructions, and then work his or her way up to more complex ones. In many situations, it makes sense to first elicit by creating sentences from simple words, then create more sentences by substitution elicitation, then create combinations of words and/or sentences, and then elicit consultant analyses. (What is meant by these methods will become clear in the subsequent sections). There is nothing wrong in doing this, as long as one understands that ultimately the resulting documentation and description will depend on a simultaneous study of all aspects of the phonology, morphology, and syntax of the language.

Also, the proportions of different data gathering methods will evolve as fieldwork progresses. Typically, there will be a lot of elicitation and little text collection at the beginning, but the proportions will be reversed – i.e. little elicitation, and a lot of text collection – at the end. But only a foolish fieldworker would ever think: "Aha, I am all done with elicitation of sentences now, time to do text collection!"

[1] As far as we can tell, the only explicit philosophy about being a descriptive fieldworker ever developed is Everett's (2004) view, derived from the philosopher William James (1842-1910), and emphasizing the concepts of coherence, empiricism, and usefulness.

[2] In contrast, the practical advice in textbooks for learning a language in the field (such as Brewster and Brewster 1976, or Healey 1975) or for studying an endangered language with the Master-Apprentice method (Hinton et al. 2002) are directional, because they are meant for language learning. Such carefully structured programs are hard to apply to data gathering.

12.1.1.3 What Data Are Worth Collecting?

Rice (2001:230), Crowley (2007:109), Bowern (2008:91) and many others urge the fieldworker to pay attention to everything the consultant says, so much so that "paying attention" may be the most important virtue of the morphosyntactic field-worker. Indeed, like astronomy, descriptive linguistics is a science where careful observation is paramount. In astronomy, we cannot travel to the stars and planets of outer space to see what they are really like; instead, we have to rely on what we can observe from light-years away. Another quotation, this one from "Return to Laughter" by Elenore Smith Bowen, a classic of anthropological fiction, speaks to this point:

> A lecture from the past reproached me: The anthropologist cannot, like the chemist or biologist, arrange controlled experiments. Like the astronomer, he can only observe. But unlike the astronomer, his mere presence produces changes in the data he is trying to observe. He himself is a disturbing influence which he must endeavor to keep to the minimum. His claim to science must therefore rest on a meticulous accuracy of observation and on a cool objective approach to his data (Bowen 1964:184–185).

The same thing is true, *mutatis mutandis*, with linguistics. We cannot get into a speaker's brain and see which neuron does what when a particular grammatical construction is used (assuming, like Chomsky, the founder of generativism, that there actually is a language organ in there somewhere). All we can do is observe what comes out of the speaker's mouth. If an astronomer observes and describes a black hole or quasar in a part of the universe, s/he can publish that observation in a scientific journal, regardless of whether or not it fits into someone's theory; people will be excited about it, and s/he will get credit for it in the academic world.

In generative formal linguistics, the situation has often been different. A linguist discovering and describing a new sort of grammatical construction will not be able to get this discovery published in a generative-oriented journal, as it will be regarded as a mere observation, not worthy of much academic credit. The construction will become important only if someone can show how it fits neatly into someone's formal theory.

For the fieldworker, it is better to observe and describe linguistic facts regardless of whether or not they fit into someone's formal theory. "Paying attention", however, does not mean that the fieldworker can or should try to pay attention to all morphosyntactic features of a language at the same time. Nida (1952–1953: 99–101) recommends "selective listening". Similarly, no astronomer pays equal attention to the whole night sky.

12.1.2 Terminology and Classification

First, some comments on the term "gathering". By now the reader will have noted our recurring usage of the term "gathering techniques". But why not talk about "elicitation", as in Swerts and Collier (1992), Senft (1995), Abbi (2001), Mithun (2001:35), Mosel (2006b:75), Crowley (2007:92), Bowern (2008:73), and pretty much any other discussion of fieldwork methodology?

One problem with the term "elicitation" is that it is vague. For some researchers, elicitation implies the use of questionnaires and questioning-without-context; it excludes data from texts. For others, elicitation potentially covers any sort of questioning methodology. Basically, the meaning of "elicitation" boils down to "data collection", or even "whatever one does to get the consultant to say something".

Another problem with the term "elicitation" is that it seems to imply data gathering without linguistic analysis. But as already stated in Chapters 1 and 2, and in Section 12.1.1.1 above, there is no such thing as data gathering separable from analysis. Some analysis in the field is unavoidable, as pointed out by Crowley (2007:151–154). Since the term "elicitation" is firmly entrenched in discussing fieldwork techniques, we retain it as shorthand for "gathering technique", with the continued understanding that "grammar gathering techniques", or, for short, "grammar elicitation" is not separable from analysis (Mosel 2006a:43).

12.2 Morphosyntax Gathering Techniques

Elicitation methods can be classified according to two criteria. The first criterion considers what controls elicitation. By definition, elicitation is controlled by the fieldworker, of course, but it can also be controlled by other factors. Samarin (1967b:112–120) distinguishes elicitation methods as either being schedule-controlled or analysis-controlled. The second criterion is the difficulty of the elicitation method.

It should be remembered that sociolinguistic data gathering techniques have been explicitly described in the sociolinguistic literature, and that *mutatis mutandis*, they are also applicable to morphosyntax. Discussions of sociolinguistic techniques are in Ammon et al. (1988), Dorian (1986, 2001), Johnstone (2000), McDavid (1985), Maynor (1982), Milroy (1987:39–67), Wolfram (1986), and Wolfram and Fasold (1974:36–72).

12.2.1 Schedule-Controlled Elicitation

This refers to what Samarin (1967b:108–112) calls "scheduled elicitation". The fieldworker has a schedule or questionnaire of material to elicit, prepared by the fieldworker or by another linguist, and asks the questions in the order of the schedule. The fieldworker's notebooks will be set up to match the prepared schedule, and the numbered answers ideally matched to the numbered questions on the schedule. If the schedule and the answers notebook are separate documents, it is also convenient to match the pagination of the answers notebook with the pagination of the schedule.

Elicitation schedules or questionnaires focusing on specific language families or language areas are usually more useful than those covering "the languages of the world".

How closely the schedule is actually followed is, to some extent, the result of negotiation between the fieldworker and the consultant. In some cases, both the consultant and fieldworker, with equal docility, will follow the schedule. More commonly, the fieldworker will find the schedule inadequate and will start deviating from it, in effect allowing the analysis to control the elicitation. And in some cases the consultant might not notice, or might be delighted by the change of pace and direction. It is also not uncommon for the consultant – particularly one with military training – to insist that the schedule be followed. There are also common cases where the consultant is so bored that s/he tries to depart from the schedule, while the fieldworker is comfortable with it. Sometimes schedules are so inadequate that the consultant sees no point in it at all. Rather than torturing the consultant with plan A, it is always advisable to have a plan B (maybe another schedule or some analysis-controlled elicitation) ready. And even a plan C.

Vaux et al. (2007) advise the fieldworker to stick to the written elicitation schedule for the day. Sometimes the point of the schedule will become apparent to the consultant, and this can affect the responses he or she gives. In this case, the fieldworker could make a note of the insight and plan to follow up on that point at a later stage, while continuing to follow the schedule (Vaux et al. 2007). But also, in our experience, common sense and the mood of the consultant and fieldworker could dictate that the schedule be abandoned for now and that the insight should be pursued right away. It is important that both the consultant and the fieldworker understand that the schedule is just a guideline; it can be modified or even set aside if need be. And even though the fieldworker is the ultimate guide, feedback from the consultant on how to best proceed with the elicitation is always welcome. Nothing impresses the consultant more than good-humored flexibility and open-mindedness on the part of the fieldworker in such matters (Newman and Ratliff 2001:6–7, Everett 2001:178–179).

There are two basic types of elicitation schedules or questionnaires, to which we now turn: Analytical questionnaires are discussed in Section 12.2.1.1, Translation questionnaires are discussed in Section 12.2.1.2, and resources for both types are provided in Section 12.2.1.3.

12.2.1.1 Analytical Questionnaires

A questionnaire asking the fieldworker analytical questions about the language is called an “analytical questionnaire”, and sometimes, rather redundantly, “question questionnaire” (König et al. 1993). Examples of questions on such a questionnaire are “Does the language have a case marking system?” If yes, “Is the case marking system nominative/accusative, or absolutive/ergative?”, and so on. These questionnaires, while useful aids to memory, depend entirely on the skill and imaginativeness of the fieldworker in eliciting the answers to the questions. Examples of such questionnaires are Dooley (1989) regarding quoted speech, van Engelenhoven (2000) regarding deixis and location, Geniušienė (1987:361–373) regarding reflexives, Klamer (2000) regarding valence, Aikhenvald (2004:385–390) regarding evidentiality, and Aikhenvald (2007:62–64) regarding word formation. A useful set of short analytical questionnaires are the sections entitled “What to investigate” at the end of the chapters in Dixon (2010b). These are questionnaires on the topics of: grammatical versus phonological word (33–34), distinguishing noun and verb (59–60), adjectives (104–108), transitivity (153), copula and verbless clauses (185–186), pronouns and demonstratives (256–257), possession (306–307), relative clauses (366–367), and complementation (415–417). It is, of course, almost always impossible to directly ask the consultant such questions. Trying to do so might produce very embarrassing results, as shown in Nida (1981:169).

The best known and longest analytical questionnaire is Comrie and Smith (1977), also called the *Lingua Descriptive Studies Questionnaire*. This questionnaire is quite long and extremely structured, and is really intended as a guide for writing extensive descriptive grammars which can be readily compared typologically. It is too long to use comfortably for elicitation in the field, but is still useful for terminology, and some of the detailed questions will provide the fieldworker with useful ideas. This questionnaire is discussed at some length by Mosel (2006a: 56–58).

The analytical questionnaires made for the South American Indian languages Documentation Project (SAILDP) were compiled under the direction of Brent Berlin, Terrence Kaufman, and Aryon Rodrigues. The project is based in the Department of Anthropology, University of California Berkeley, and includes a 33-page phonological and morphosyntactic questionnaire, based in part on Comrie and Smith (1977) (Kaufman 1990:28).

Three other collections of analytical questionnaires focusing on the elicitation of particular morphosyntactic phenomena are worth mentioning. One set is available through Eurotyp, a program in language typology from the European Science Foundation (König et al. 1993). An example is the Eurotyp Word Order Questionnaire by Siewierska (1993), which contains two pages of explanations (pp. 1–2), and twenty-two pages of questions (pp. 3–25).[3]

Another collection (Max Planck Institute for Evolutionary Anthropology 2007) makes available online most of the Eurotyp questionnaires, the Comrie and Smith (1977) questionnaire, and several others.

There is also a collection of manuals with questionnaires published by the Max Planck Institute for Psycholinguistics in Nijmegen. An example is the Demonstratives questionnaire by Wilkins (1999), supplemented by Enfield and Dunn (2001).

Because analytical questionnaires pose such broad theoretical questions, they are not easy to use for beginners, and they can sometimes be a frustrating straight-jacket for advanced fieldworkers.

12.2.1.2 Translation Questionnaires

The most widespread type of questionnaire is the questionnaire with lists of sentences to translate into the target language. It is called a "translation questionnaire", an "elicitation questionnaire", or a "primary data questionnaire" (König et al. 1993).

[3] It also includes about 40 sentences to be considered as examples or to be translated, which according to the Eurotyp guidelines (König et al. 1993) is not enough to have it qualify as a Mixed questionnaire.

The term "elicitation questionnaire" is somewhat confusing, since we already use the term "elicitation schedule" for any type of questionnaire. We will therefore call them "translation questionnaires", although Eurotyp prefers the term "elicitation questionnaire" (König et al. 1993).

One problem with most such questionnaires is that they contain sentences out of context (Chelliah 2001). This questionnaire type, while seemingly the most reassuring and easiest to use for the beginning fieldworker, is basically the scheduled variety of the Reverse Translation elicitation method, to be discussed below. As we will see, it is downright dangerous for syntactic elicitation.

A third type of elicitation schedule is theoretically possible. This would be a "metaschedule", drawing upon specific elicitation methods (other than Reverse translation) – such as Word-list-based elicitation, Paradigmatic elicitation, Syntagmatic elicitation, or Transformational elicitation[4] – for specific grammatical topics or constructions. This would be very useful for beginners, but unfortunately no one has attempted to systematically develop this idea. The helpful hints on how to elicit certain types of information found throughout Healey (1975), Abbi (2001), and in Bowern (2008:77–78) are perhaps the closest we have to such a metaschedule.

12.2.1.3 Survey of Analytical and Translation Questionnaires

Most grammatical questionnaires remain unpublished, and they are not nearly as numerous as those that exist for lexicographical elicitation or for dialectological work. Questionnaires that mix the analytical and translation types can be called "Mixed", and some Eurotyp questionnaires are of this sort (König et al. 1993). An early survey of linguistic questionnaires is in Pop (1955); a list of references to questionnaires supplementing Pop is in Healey (1964:22–23). Most of the lexicographical elicitation or dialectological elicitation manuals contain lists of sentences to translate, so the caveats regarding Reverse translation elicitation (discussed in Section 12.2.2.6, below) apply.

Table 12.1 is a sampling of general grammatical questionnaires compiled since 1950.[5] The table shows: a reference; the language family or area covered; page numbers where suggestions for use can be found; whether the questionnaire is analytical or translation, and, if it is a translation questionnaire, page numbers and number of grammatical items to translate; page numbers and number of lexical items to translate[6]; and additional comments.

Several of the questionnaires above, such as Sutton and Walsh (1979) and Bouquiaux et al. (1992) were compiled based on other, usually unpublished, questionnaires for the same areas.

[4] These elicitation methods are described below under Easy methods.

[5] This is a sampling of questionnaires we have examined; it does not include others such as Johnston (1980) on Oceanic (Austronesian) languages, which we were unable to obtain.

[6] Questionnaires designed for lexical elicitation only are discussed in Chapter 8.

Table 12.1 A sampling of general grammatical questionnaires

Reference (chronological)	Language family/ area covered	Author's guidelines for usage given on:	Analytical or Translation	Lexical items to translate	Comments
Cohen (1950–1951)	Any area or family	pp. 13–15	Translation: 45 items	231 items	Trilingual French-English-Spanish
Whiteley (1964)	Bantu languages of Africa	pp. 5–14	Analytical	pp. 16–19; 662 items	
Bridgeman and Gudschinsky (1965)	Tupi languages of Brazil	pp. 23–25	Analytical	None	In Portuguese
Summer Institute of Linguistics (1965)	Brazilian indigenous languages	pp. 27–32	Translation: pp. 33–36	Mixed in with sentences to translate	In Portuguese
Capell (1965)	Australian languages	pp. 1–9	Translation: 144 short sentences	446 items	Organized by semantic field
Freeze (1989) taken as an example for the Archivo de Lenguas Indígenas de México (1969–present) format	Mexican indigenous languages	pp. 16–17	Translation: pp. 73–151; 594 items	pp. 153–168; ca. 532 items	In Spanish
Prost (1974)	West African	pp. 75–89	Analytical	None	In French
Tucker (1974)	Bantu and non-Bantu of Africa	Interspersed with questionnaires	Translation: pp. 60–73; ca. 560 items	Mixed in with sentences to translate	In French
Healey (1975)	Papua New Guinea	Passim	Translation: pp. 460–464; ca. 185 items	pp. 452–459; ca. 450 items	
Comrie and Smith (1977)	Any area or family	pp. 5–9	Analytical	pp. 66–71; 207 items	(Discussed above)
Kibrik (1977)	Russia, Caucasus	pp. 28–35	Translation: pp. 83–100; ca. 385 items	pp. 103–124; 1,482 items	Translated from Russian
Sutton and Walsh (1979)	Australian aboriginal languages	pp. 23–33	Analytical	pp. 37–53; ca. 2,195 items	

(continued)

Table 12.1 (continued)

Reference (chronological)	Language family/ area covered	Author's guidelines for usage given on:	Analytical or Translation	Lexical items to translate	Comments
Berlin and Kaufman (1985–1987)	South American Indian languages	pp. 1–26	Analytical	Includes a separate Lexical questionnaire of 181 pp. and ca. 1,800 items	Trilingual Spanish-Portuguese-English; (discussed above)
Lastra de Suárez (1986)	Nahuatl languages	pp. 10–13; grammatical index pp. 253–245	Translation: pp. 245–250; ca. 167 items	pp. 239–245; ca. 264 items	In Spanish, designed for dialectology
Bouquiaux et al. (1992)	African languages	Interspersed with questionnaires	Translation: pp. 257–367; nine morphosyntactic questionnaires; the longest one has 6,406 items	pp. 186–193, pp. 201–229; ca. 872 items in sentences; eight other thematic questionnaires	Translated from French
König et al. (1993)	Any area or family	Chapter 5, p. 4	References provided to 32 questionnaires: 17 analytical, 4 elicitation, and 11 mixed	None	
Abbi (2001)	South Asian languages	pp. 68–72	Translation: pp. 248–262; 326 items	pp. 244–247; 407 items	Hindi equivalents given occasionally; 55 Hindi-English sentences

12.2.2 Analysis-Controlled Elicitation

Also called “analytical elicitation” by Samarin (1967b:112), under this type of elicitation there is no schedule. The analysis (in the most informal sense of the word) controls the elicitation. Analytical elicitation is more daunting for beginners, who often wonder where to start. Good advice regarding this is in Dimmendaal (2001:66), Everett (2001:179–182), Hale (2001), Crowley (2007:95–102), Vaux et al. (2007:205–208), and Bowern (2008:34–37, 73–77). These sources suggest that no one really starts with analytical elicitation at the very beginning of fieldwork. Probably every fieldworker starts with some sort of scheduled elicitation to establish some of the basics of phonetics and phonology (see Chapter 10), and to gather a basic vocabulary (see Chapter 9). However, to make serious progress in the realms of morphology and syntax, it is important to downplay schedules and to start relying more heavily on analytical techniques.

Hale (2001:84–85) points out that, in any analytical elicitation – but particularly in syntactic elicitation – it is essential to have some sort of “script” (or “protocol”): in effect, a very informal schedule which serves to remind the fieldworker of what s/he is looking for and where s/he is going. This will simplify organization of the fieldworker’s notes, and will also give the fieldworker something to fall back on when s/he runs out of analytical ideas. Bowern (2008:80–81, 90) and Crowley (2007:95–96) also advise designing one’s own protocols or questionnaires.

In choosing an approach to analysis-controlled elicitation, it is important to take into consideration the degree of obsolescence or endangerment of the language. As shown in Mithun (1990), there are many sorts of analysis-controlled elicitation that cannot be used effectively when speakers are no longer dominant in the target language. Ancillary (or text-based) elicitation becomes the best method, and sometimes the only reliable one. And when documentation of what people actually say in seriously endangered languages is a priority, any sort of elicitation based on constructed sentences should be given lower priority, as demonstrated with several languages by Mithun (2001, 2007a, b), and text-based elicitation should be given the highest priority.

We classify the different types of analysis-controlled elicitation according to the sorts of prompts or stimuli they utilize. We distinguish: Target language interrogation elicitation (Section 12.2.2.1); Stimulus-driven elicitation (Section 12.2.2.2); Target language manipulation elicitation (Section 12.2.2.3); Target language translation elicitation (Section 12.2.2.4); Target language construction and introspective judgment elicitation (Section 12.2.2.5); and Reverse translation elicitation (Section 12.2.2.6).

We also discuss three analytical tasks which are essentially combinations of several of these approaches. These are: Review elicitation (Section 12.2.2.7), Ancillary elicitation (Section 12.2.2.8), and Covert elicitation (Section 12.2.2.9). Finally, we discuss the elicitation of consultant analyses, which, since it is the elicitation of an analysis, will be called Meta-elicitation (Section 12.2.2.10).

We also rate these elicitation tasks in terms of difficulty, with an annotation in brackets next to each type. By (easy) we mean tasks that are likely to give the fieldworker reliable information without too much strain on the fieldworker/consultant collaboration. By (medium difficulty) and (difficult), we mean tasks that are likely to give the fieldworker reliable information only by putting some strain on the fieldworker/consultant collaboration. By (perilous), we mean tasks that not only put a strain on the fieldworker/consultant collaboration, but are likely to produce unreliable results as well.

Needless to say, beginning fieldworkers will want to try the easy tasks first, but since there is no directionality requiring one to use easy tasks before the difficult tasks, s/he will be led onto a garden path of difficult and perilous tasks faster than s/he thinks. The main point is to remain aware and on one's guard.

12.2.2.1 Target Language Interrogation Elicitation (Easy)

This is elicitation where questions are asked in the target language. Questions like "What is it?", "Where is it?" and "What are you doing?" are helpful to the fieldworker in developing conversational fluency in the language, and of course will be used in monolingual situations (Everett 2001). One should always record not only an answer specific to the situation, but a variety of appropriate answers as well. This tells the fieldworker about the real meaning of the question (Healey 1964:13). A more complicated type of interrogation is to describe a situation and ask what the person might say in that situation at a given moment (Mosel 2006b:76). For example: "Tell me what s/he is doing?" It is also possible to act things out oneself – for instance, "Tell me what I am doing" – but that is obviously less natural. Advanced fieldworkers might ask "Tell me how you would ask X for a cup of coffee", in situations where this is not obvious.

12.2.2.2 Stimulus-Driven Elicitation

In this type of elicitation, the consultant is provided with some concrete objects, or pictures, video clips, or a movie, and is asked to comment on them. Several subtypes are distinguished here.

Prop-Driven Elicitation (Easy)

Consultants generally enjoy this type of activity. For number marking, for example, the consultant is presented with one object, then two, then three, and so on. Some examples of props are in Everett (2001:179–180).

Prop-driven elicitation is particularly useful for the study of noun classification (Section 11.2.3.2.1). The fieldworker will need a number of different types of objects varying in animacy, shape, size, surface, consistency, and paired or not paired, as well as objects contained in different ways: in a plate, cup, basket, bottle, or bag. It is important to realize that noun classification is bound to interact and/or fuse with other noun marking (e.g. gender or noun class, case, and number). (Noun classification can also be analyzed through texts, which often provide unusual and unexpected nouns or categorizations, as discussed in the section on Ancillary elicitation, below). Prop-driven elicitation is also useful for the study of locative expressions (Levinson 1992, Hellwig 2006:330–331). The fieldworker should remember that his/her own body is one big prop, and pointing at body parts is prop-driven elicitation.

The main limitations of prop-driven elicitation are the following:

One can only talk about objects within sight, ideally objects on the table in front of the fieldworker and consultant. For other things, pictorial stimulus-driven elicitation can be used.

Of course, the deictic nature of props needs to be taken into account. So the sentence that the fieldworker thinks means ‘I pick up this knife’, might well mean ‘you pick up that knife’.

Finally, and most important, as soon as something slightly abstract is pointed at, misunderstandings will occur. In some cases, a fieldworker has tried to get the term for “mirror” by looking into the palm of one’s hand, and to get as translation something like ‘you are looking at it’, or ‘your hand’.

But on the other hand, there are a few rather abstract concepts that are best demonstrated by gestures; as Haviland (2006:144–145) points out, it is universally hard to explain what a spiral is without using gestures.

Pictorial Stimulus-Driven Elicitation (Medium Difficulty)

This involves the use of pictorial stimuli, including line drawings, photographs, and video-clips. Driven elicitation has been used mostly for the elicitation of single lexical items in dialectology, and pictorial manuals for dialectological elicitation such as Sapon (1957) are available, but this method can be used to elicit morphological and syntactic information as well. In this method, the fieldworker shows a picture to the consultant and asks him/her to describe it, or to comment on it (Harris and Voegelin 1953, Hayes 1954, Aitken 1955, Yegerlehner 1955).

Sentences containing quantifiers such as *every*, *some*, *each*, and *all*, as in *Every man kisses some woman* are notoriously difficult to elicit through translations (Matthewson 2004:384). In a groundbreaking study of distributive numbers, Gil (1982) used pictures of people carrying boxes and suitcases in various combinations. In some languages, however, the verb “to carry” varies depending on the shape of the box or suitcase, which might cause translation difficulties for consultants speaking these languages (Gil 2001:103, Munro 2003:138).

Pictures, picture games, and short video-clips are also useful for the study of spatial or locative relations, as described in Levinson (1992), Vandeloise (1991), Hellwig (2006:332–340); for property concepts (Hellwig 2007); and for ideophone syntax (Nuckolls 1996).

For research on verbs of motion and change, O'Connor (2007:18) refers to a list of useful stimulus tasks from the Max Planck Institute for Psycholinguistics, including line drawings, photographs, photo matching games, and video-clips.

For verbs of action, acting out scenarios is more useful than pictures in getting the meaning across. When several verbs of action are being elicited, and the results need to be easily comparable, video-clips of someone acting them out can be a useful stimulus (Bohnemeyer et al. 2001).

In general, language textbooks or phrasebooks that teach through pictures are useful, too. A wordless picture book for children that has been used quite a bit for eliciting discourse is "Frog, Where are You?" (Mayer 1969). Hayes (1954) is a detailed account of picture elicitation with stick figures from the "Through Pictures" collection. It is best to use materials that do not have any legends that the consultant can read. Frances Ingemann (p.c.) liked to use the textbook "Hebrew Through Pictures" (Richards et al. 1961) in work on a Papua New Guinea language, because the accompanying sentences were written in the Hebrew script, and so were not a distraction to the consultants. A good example of sentences gathered through picture elicitation is in Seiler (1970).

Some of the pitfalls associated with picture elicitation are the following:

As with prop-driven elicitation, the fieldworker needs to remain aware of the deictic nature of visual stimuli (Wunderlich 1981). The consultant might respond from the perspective of the speaker, and not from the perspective of the objects. So if the consultant thinks that the man in the pictures could be himself, he might well comment: "I am in front of the tree", rather than "The man in the picture is in front of the tree" (Hopkins and Furbee 1991:66–73).

If there are too many pictures on one sheet, or if there are discrepancies in the size of the depictions (for example a series of animals of different actual sizes, but all presented as the same size – as occurs in some field guides to animals or plants) – consultants will be easily confused.

Finally, some consultants simply do not like to comment on pictures or video-clips. Cultural issues arising from the use of picture or video-clip descriptions are addressed in Enfield (2002), in Crowley (2007:108), and in Bowern (2008:82–84).

A short silent movie such as the Pear Story (Chafe 1980), is of course primarily used to elicit discourse (Chapter 13), but if several speakers watch it and retell the story, it can be quite useful for gathering of morphosyntactic detail as well. The Pear Story movie can also be profitably used to elicit a running commentary, with the consultant describing the events of the film as he or she is watching it.

12.2.2.3 Target Language Manipulation Elicitation

In this type of elicitation, some word or structure of the target language is manipulated, and the consultant is asked to react to it, or correct it.

While working with this method (as well as with Target language construction and introspective judgment elicitation (Section 12.2.2.5)), the fieldworker will note that some consultant reactions or corrections are more difficult to interpret than others.

Some consultants are what we irreverently call “Pragmatics Police”. They reject anything that is ever so slightly pragmatically odd. Pragmatic police generally get into arguments with people who are more imaginative pragmatically. Needless to say, the fieldworker must stay out of such arguments.

Some consultants are what we irreverently call “Culture Police”. They reject anything that is ever so slightly culturally inappropriate. An example of a sentence rejected by such people can be “I love you.” In a culture where love is not expressed in explicit words, this sentence would indeed be very inappropriate. In some cultures, comparison constructions are hard to elicit, because the culture does not emphasize comparison of things or people (Dimmendaal 2001:70). Other examples are in Crowley (2007:99–100). Culture Police also get into arguments with people who are more imaginative in abstracting from their own culture. Again, the fieldworker must stay out of such arguments.

Several types of target language manipulation elicitation can be distinguished.

Word-List Based Elicitation (Easy)

In this method, the fieldworker asks the consultant to make up a sentence using the word “X” (Mosel 2006b:77). The words used are carefully selected from a previously collected word list (Chapter 9), so the fieldworker can get samples of different types of sentences. For example, the fieldworker might choose typically intransitive or transitive verbs as prompts. Sentences collected in this way impose minimum control over consultant production, allowing the consultant to shape the corpus. This task is so transparent that when it has been used just a few times, the consultant naturally learns it as an easy tool for language study (Hale 1965). As with all translation tasks, the translations will be fine-tuned as the fieldworker learns more about the target language. One of the drawbacks of this method is that the consultant might hit upon a sentence frame that works most of the time, so that for “pen”, “book”, “plate”, and “glass”, s/he might say: “the pen is on the table”, “the book is on the table”, “the plate is on the table”, “the glass is on the table”, and so on. This redundancy does not provide a great deal of morphosyntactic variation. Then again, for many languages, the verb “to be (at a location)”, will vary depending on the shape or animacy of the item considered, so it might be worthwhile for the fieldworker to persevere a bit before asking the consultant to diverge from the “the ---- is on the table” frame.

Paradigmatic Substitution Elicitation (Easy)

Paradigmatic substitution elicitation – also known as Frame and substitution, Substitution, or Contrastive elicitation – works best with a language which both fieldworker and consultant can write. The fieldworker writes a sentence on a blackboard or on paper. Then s/he deletes a word or constituent from the sentence, and asks for another word or constituent that can fit that slot. Asking for another word or constituent is preferable to suggesting one because it avoids the creation of artificial forms by the fieldworker (Hale 1965, Samarin 1967b:115–117, Kibrik 1977:58).

Fieldworker-Driven Paradigmatic Substitution Elicitation (Medium difficulty)

This method of elicitation is like Paradigmatic substitution elicitation except that the open slot is filled with another word or constituent by the fieldworker him/herself. Here the danger is of of course that the fieldworker might be suggesting an artificial form. The fieldworker then asks for the meaning of the resulting sentence. One also has to be aware that the one word or morpheme that the fieldworker replaced might have more than one meaning.

Paradigmatic substitution elicitation is nevertheless useful for hard-to-define elements like evidentials, or for the types of sentence adverbs found in Russian or German called “modal particles” (Abraham 1991a, b). The fieldworker will notice such elements in conversations and texts, and they may occur in elicited sentences – but that is not a sure thing. Rather than asking for the meaning of the such sentences, the method of Paradigmatic substitution elicitation can be used, and the fieldworker can then ask the consultant to explain the context in which each sentence would be appropriate (Kibrik 1977:58).

For example, to elicit paradigms, the fieldworker should begin with a simple sentence containing the word s/he wants to study. Vaux et al. (2007:233–234) suggest a sequence of clauses such as those in (1), where just one or two elements of the sentence are changed at a time in order to keep close control over the data:

(1) The woman left yesterday.
You left yesterday.
We left yesterday.
He left yesterday.
Raja left yesterday.
Raja will leave tomorrow.

Syntagmatic Elicitation (Easy)

Syntagmatic elicitation – also known as Appropriate adjustments elicitation – can be used, among other things, to see how tense is indicated. For instance, the fieldworker might ask the consultant to insert a temporal adverbial into a sentence. Beginning with the original sentence: *John went to camp*, the adverbial *tomorrow* is inserted: **John went to camp tomorrow*. What needs to be changed to make it grammatical? John *will go* to camp tomorrow. To check for person, the fieldworker could ask the consultant to change the person of the noun, and then note what adjustments are made. This procedure can be used for number, gender, and many other grammatical categories (Hale 1965). Typological knowledge of the possibilities of likely "appropriate adjustments" is crucial in this sort of elicitation. Too many adjustments – or bizarre adjustments – will confuse the consultant, or will lead him or her to begin to doubt the mental faculties of the fieldworker.

Transformational Elicitation (Medium Difficulty)

The fieldworker might ask the consultant to provide question or negative versions of sentences s/he has created. Hale (1965) points out that these are abstract, but some consultants learn to do this quickly. One difficulty with transformational elicitation is that the fieldworker and the consultant will have to agree on a transformational metalanguage (Samarin 1967b:203–204). A useful account of transformational elicitation in Vietnamese by Richard Pittman is in Healey (1975:389–398).

Fieldworker-Driven Transformational Elicitation (Difficult)

In this method, the fieldworker begins with a sentence in the target language and modifies it in some way, such as moving a constituent. Then the fieldworker asks for a grammaticality judgment, or asks about the pragmatic situation where the "changed" sentence can be used, and asks about any socio-linguistic details. This method is called Data manipulation in Bowern (2008:81–82). It can be problematic if the resulting sentences sound too contrived. Also, speakers may have a hard time judging the grammaticality of the resultant constructions (Abbi 2001).

Corrective Elicitation (Difficult)

Also called Stimulus interrogation, this method involves the fieldworker deliberately producing an ungrammatical sentence or using an incorrect form to test a hypothesis (Samarin 1967b:117). The fieldworker produces the form and then waits for the consultant's reaction. This method is useful in gauging how well the speaker is able to correct the fieldworker. Is the consultant accommodating, uncaring, or afraid to hurt the fieldworkers' feelings? When planning to use this method, the fieldworker might want to begin the fieldwork session with a request that the consultant be a strict teacher.

12.2.2.4 Target Language Translation Elicitation (Medium Difficulty)

In this method, the fieldworker asks the consultant to translate target language text or other materials gathered during previous sessions. Such translation work, while not particularly easy for the fieldworker, is often very hard for the consultant. The translations will likely be relatively free, since a consultant's word-for-word translations might be difficult for the fieldworker to understand or interpret correctly.

Some consultants are very skillful translators, some can offer only a brief or inaccurate synopsis of what the text is about, some can only do word-for-word translations, and others – though perfectly fluent in both the target language and the language the fieldworker uses for translation – are simply incapable of translation. In order to protect the consultant from embarrassment, the fieldworker needs to know the consultant well enough to be able to judge his/her abilities before embarking on a translation project. Clearly, if the consultant has trouble translating short sentences, s/he will have trouble translating discourse. On the other hand, if the consultant has trouble with short sentences out of context, s/he might well turn out to be more skillful with discourse in context.

It is also good to remember than any translation is negotiated. The speaker cannot help but read cues from the fieldworker on how s/he is performing as a translator, and the fieldworker cannot help but guide the translation in almost unconscious ways.

Matthewson (2004:388) points out that the consultant might offer ungrammatical English translations to indicate that the equivalent in the language is also ungrammatical.

Translation is quite useless in some areas of morphosyntax. Here are three examples: In deictic-rich languages such as Hän Athabascan of Alaska, or Eskimo (Woodbury 1993), the consultant will translate half a dozen deictics extremely vaguely as: "over there", another half dozen as "up there", and another half dozen as "down there", and so on; with translation, the distinctions between them are hard to identify. For the description of spatial deixis, special techniques with pictures and schedules need to be used (Levinson 1992). Passives are also hard to translate. The intransitive version of a sentence and the passive version may translate the same way. Negative questions are a third area where translation is very unreliable.

Finally, the accuracy and reliability of translations can be affected by differences in the variety of English (or any other contact language) spoken by the fieldworker and the consultant. Fieldworkers who are speakers of "non-standard" dialects of English or who are non-native speakers of English are likely to misunderstand some of the consultants' translations into English. This happens because the fieldworker is so focused on the target language that s/he might forget to notice dialectal differences in the contact language; inaccurate translations can result. Chelliah (2001:157–158) has useful examples of problems where the contact language of the fieldworker is American English, but the contact language of the speaker is Indian English.

12.2.2.5 Target Language Construction and Introspective Judgment Elicitation (Perilous)

In this type of elicitation, the consultant goes beyond Target language manipulation (Section 12.2.2.3) and constructs sentences in the target language. But rather than asking the consultant to correct them, or to comment on them extensively, the fieldworker simply relies on the consultants' intuition and their ability to introspect, and asks them to judge if the constructed sentence is grammatical or not. This is, in simple terms, the "can you say this?" method, and the fieldworker is content with a simple "yes" or "no" answer. The method is fast, and appears to be easy on both the fieldworker and the consultant.

With the advent of Chomskyan linguistics and the importance of introspection and reliance on native speaker intuition, it came to be expected that fieldworkers would rely on their consultants for grammaticality judgments. Many fieldworkers, particularly in syntactic fieldwork, assume that elicitation of grammaticality judgments from consultants on constructed data is a reliable approach. This is untrue for quite a few reasons.

First and foremost, regardless of what Chomsky (1957) claimed, and regardless of the extensive defense of introspection by Newmeyer (1983:48–66), the insights and intuitions of native speakers are simply not reliable. There is plenty of literature (Labov 1972, 1973, 1975, 1996, Ross 1979, Hinds 1981, Ulvestad 1981, Sampson 2002) demonstrating that people do not know what they do and do not say, and that grammaticality judgments vary not only across idiolects of different speakers, but even for a single speaker at different points in time. Judgments become blurry. More common-sense advice regarding grammaticality judgments is in Crowley (2007:102–104) and in Bowern (2008:78–80).

One of the earlier reactions to Chomsky's claim was a series of experiments conducted by Maclay and Sleator (1960) on judgments of grammaticality. One disturbing result was the following: three out of twenty-one of their subjects judged the utterance *Label break to calmed about and* to be grammatical (Maclay and Sleator 1960:282). The researchers concluded that a way out was to gather a statistically significant number of responses to check the validity of responses so that the analyses could be regarded as valid. However, while it is certainly possible to take such a statistical approach to the analysis of grammaticality judgments, it was never Chomsky's intention to base his theory of universal grammar on such methods. And whether fieldworkers should base their grammatical analyses on statistical methods is a matter of debate. Let us just say here that, from a practical point of view, fieldworkers cannot always gather enough data or work with a wide enough variety of consultants to seriously claim that their data is statistically significant.

In addition to the fact that introspection is unreliable, there are a variety of other reasons why consultants cannot provide grammaticality judgments reliably. We list them here.

Naturally enough, consultants aim to please the investigator, and when pressed to provide numerous judgments on utterances – usually given in isolation, with no context, or in a weird context – they might settle into a sort of psychologically dulling automatic pilot; having gotten an idea of what the investigator likes, and what s/he does not want to hear, they unconsciously (sometimes consciously) develop a feel for which ones they should say "yes" to and which ones they should reject (Sutton and Walsh 1979:32). It is a sort of dressage effect, like the famous horse Clever Hans.[7] Consultants figure out very quickly that the investigator does

[7] The story of the horse Clever Hans can be easily found online, e.g. in http://en.wikipedia.org/wiki/Clever_Hans.

not like it if they say “yes” to every utterance; they also know that the fieldworker likes it even less if they say “no” (or worse, look outraged) at every utterance (Nida 1981:173). So the consultant develops a feel for what the fieldworker can tolerate, and what s/he cannot. An impression of “working well together” is then created, and is cozily shared by the consultant and fieldworker.

Another problem with basing analysis on native-speaker grammaticality judgments is that consultants with knowledge of a grammatical tradition may reject everything that is not prescriptively correct. Or relatively grammatically sophisticated consultants may err on the side of caution and accept too much as correct. My [de Reuse’s] grammatically sophisticated Western Apache consultant would say something like, “Well, I have never said this, and I have never heard this, but what you just said sounds good. Good job, Willem!” It is better to allow the consultant to be creative, rather than let him/her praise the fieldworker for his/her own creativity!

Some consultants are quite forgiving regarding the fieldworker’s pronunciation, while others will reject anything not produced with native or near-native pronunciation.

A consultant might reject a sentence because it is factually inaccurate, or s/he might reject a sentence because it is factually accurate, if the truth revealed by it is embarrassing to him/her. One should also keep in mind what was said about “Pragmatic Police” and “Culture Police” in Section 12.2.2.3.

Training one’s consultant in linguistics is a laudable educational pursuit, but as far as improving the reliability of introspection is concerned, it does not seem to be of much help. As Anttila (1972:349) points out, linguistic training spoils the intuitions of the native speaker, and native speaker grammarians are no longer normal speakers. Further discussion of this point is in Spencer (1973). And the fieldworker cannot rely on his/her own judgments even if s/he becomes as fluent as a near-native speaker in the target language.

Now, having said all this, it would be incorrect to say that introspection by the native speaker is useless. Sometimes, introspection by native speakers yields reliable and valuable results. One case in point is the semantics of two Lakota (Siouan) postpositions, which Rood (2003) managed to elucidate by simply asking about the meaning difference. Earlier extensive investigation of texts had suggested that these postpositions were synonyms.

Our point is that introspection regarding grammaticality judgments of constructed sentences is far from infallible. The most careful discussion of the strengths and weaknesses of introspection is in Talmy (n.d.), who compiled a detailed list of what introspection is and is not reliable for. Other discussions of intuition and introspective judgments are in Carroll et al. (1981), Dale (1978), García (1967), Ringen (1977), Schneider (1995), and Schwarte (1974). The relationship between grammaticality judgments and the consultant’s native speaker status is discussed in Coulmas (1981b), Itkonen (1981), Paikeday (1985:28, 41, 51–53, 56, 78, 101), Ringen (1981), and Schnelle (1981). We discuss the issue of grammaticality judgments regarding syntax in Section 12.4.2.

12.2.2.6 Reverse Translation Elicitation (Perilous)

Simply asking one's consultant to translate sentences from the contact language into the target language seems like a very obvious and easy way of getting data. One of the most common questions fieldworkers ask is, indeed: "How do you say *X* in your language?". This is called Reverse Translation Elicitation (Samarin 1967b:114–115).

The problem is that most consultants cannot do this reliably, and will translate word for word. Even if one is lucky enough to work with a skilled translator, the results can still be odd, unnatural, and/or unidiomatic. The worst problem is that the fieldworker will have no way of knowing which translations are natural and idiomatic, and which ones are not. The result is a sentence grammar straight from English, and ultimately the grammar written on the basis of these data will be, in Mary Haas' words (p.c. to de Reuse), "a grammar of translation into English". Or, to put it in Everett's (2001:185) words, reverse translation affects the "purity of data collected".

Dixon (2007), Mary Haas (p.c. to de Reuse), and Munro (2003:145) advise strongly against this method, with good reason. So did Bloomfield. According to Voegelin (1959a:114–115, 1960:204), Bloomfield preferred to learn the informant's language, and use that in order to get his informants to talk; he would check his data by making child-like errors and see what kinds of corrections were offered, without asking a "how do you say...?" question.

This does not mean that reverse translation should never be used, but just that extreme caution is advised.[8] Voegelin (1959b:210–213), for example, was less dogmatically opposed to reverse translation himself; he used the general term "text eliciting", which includes reverse translation, and then proposed "shared morpheme elicitation", where the translation into the contact language is modified a bit, and the fieldworker resorts to reverse translation to elicit a new sentence. For example, if the fieldworker has gotten "Then she is talking to me", s/he asks for: "Then she is talking to them." This, of course, depends on the fieldworker's knowledge about what could be a shared morpheme in the target language. Also, the longer the sentence, the more unreliable "shared morpheme elicitation" becomes.

If the consultant is literate, it is possible to give him/her a written task. That way, the consultant has time to think over a written translation; written translations tend to be a bit more reliable (although often unidiomatic or overly formal) than immediate oral translation. It might be worthwhile to find another consultant to retranslate the sentences back into the contact language, and see what happens.

Another way of using reverse translation is to use very short colloquial sentences, such as "where are you going?" and responses such as "I am going to X". It is easier for the consultant to give accurate translations for these. Some pedagogical books with natural-sounding conversational sentences are Kallionen

[8] Interestingly, Samarin (1967b:54) suggests that confirmation of Voegelin's statements about Bloomfield's method is needed.

(1974, Finnish conversations), Mintz (1986, a textbook of Malay), Palmer (1974, a textbook of Thai), and Toulouse (1995, a phrasebook of Ojibwe (Algonquian)). My [de Reuse's] experience with short colloquial sentences involved translations from English to Siberian Yupik Eskimo. The English colloquial sentences were taken from quite reliable pedagogical materials published on Central Yup'ik Eskimo (Hensel et al. 1983), a closely related language. Consultants had no trouble doing this, but over the years I have heard reports from other Siberian Yupik speakers that there is something unnatural about some of those sentences. This unnaturalness is due in part to the fact that pedagogical materials – no matter how conversational – are always somewhat artificial, but I think it is due largely to my use of the English translations to elicit Siberian Yupik Eskimo equivalents.

I [de Reuse] also elicited quite a number of Western Apache (Athabascan) data on the basis of Navajo conversational sentences from textbooks such as Wilson (1969), and Wilson and Dennison (1978). The languages and cultures are very close. I have not heard any complaints from Apaches yet, because my consultants are polite people, but sooner or later, I expect to hear some. Whatever one does, reverse translation just never works very well. More discussion of reverse translation is in Bowern (2008:74, 78, 85).

Because reverse translation is so difficult, one good approach might be to get translations in the morning, and then meet with a group of speakers in the evening to discuss them. In such a discussion, speakers might help weed out the problematic sentences.

12.2.2.7 Review Elicitation (Medium Difficulty to Perilous)

Also called Cross-interrogation elicitation, this procedure is used to check facts already elicited, and therefore presupposes some method of previous elicitation. The fieldworker reviews previous materials together with the consultant, and looks for agreements, disagreements, variants and "mistakes". As a controlling device, it is useful to ask for a translation back into the target language of something previously elicited (Dimmendaal 2001:70–71, Bowern 2008:76).

If the consultant finds a mistake, s/he may not always be able to say why it is wrong, or even offer a correction (Nida 1950). The fieldworker should remain non-confrontational. S/he might say that s/he wrote it down wrong, etc. The consultant is never wrong – even though, as we all know, the consultant can be wrong for a variety of reasons (Crowley 2007:132–134). After working with the same speaker for more than 20 years, I [de Reuse] now feel more comfortable in overtly pointing out a discrepancy and asking for clarification, but I can only do that with friends.

It is important to check and recheck the data with as many speakers as possible. But at the same time, it is not necessary for all the data one collects to be rechecked. The fieldworker develops a feel for which data or consultants are most reliable, and keeps these feelings to him/herself. Checking for data inconsistencies by comparing across several sessions is also best carried out by oneself.

12.2.2.8 Ancillary Elicitation (Easy to Difficult)

This sort of elicitation, first named by Voegelin (1959b:218), is also called Text-based elicitation (Samarin 1967b:117–119). Texts are crucial for morphosyntactic elicitation (Dimmendaal 2001:71–72). As Nida (1947:146) points out:

> Only texts (...) can serve as the basic core of data from which an adequate analysis can be made.

An extensive discussion of ancillary elicitation – with numerous examples – is in Chelliah (2001). In this approach, the fieldworker departs from a previously recorded and translated text (Section 12.2.2.4), ideally a text of naturally occurring discourse. S/he then elicits more information on the sentences of that text by using the methods described in Section 12.2.2.3 (Nida 1981:170). It is best to start working with shorter sentences. For example, based on a text, one could develop a hypothesis about word order and constituent structure. The fieldworker could then change word order to see if it is strict or variable.

Consultants might have completely different expectations of what the fieldworker is expected to understand or not understand in a text. Loanwords from English, which sound assimilated to the fieldworker, might be considered unassimilated by the consultant and therefore obvious to everyone. This is what I [de Reuse] like to call the "[pɪna:naməŋ] problem". I was once going over a text with a consultant of Central Siberian Yupik Eskimo, and asked him what the word [pɪna:naməŋ] was. The consultant gave me an exasperated look (implying: "how can you NOT know that"?), and said, "well, that's the banana". I had not recognized the word "banana" with *-məŋ*, the modalis singular case ending![9]

It is not unusual to come across speakers who cannot manipulate a text in any way. Needless to say, these speakers should not be relied upon for ancillary elicitation, but they might be helpful with other tasks.

Ideally one would want to use ancillary or text-based elicitation as much as possible, and reverse translation (Section 12.2.2.6) as little as possible. But what if the fieldworker wants to study relative clauses, and there turn out to be very few relative clauses in a text? In such a case some cautious reverse translation might be attempted, but the results of this sort of work should be kept under the fieldworker's hat, as it were, until s/he can find textual or conversational corroborations for these putative relative clause constructions.

12.2.2.9 Covert Elicitation (Easy to Perilous)

In this methodology, it appears that a particular elicitation procedure is being followed (any of those described in Sections 12.2.2.1–12.2.2.8), but in fact, the

[9] If the reader is curious why we were discussing a banana in Yupik, I [de Reuse] had asked the speaker to comment on a story in a children's book. I had apparently forgotten there was a banana in the story.

fieldworker is actually paying attention to something else, and the consultant is not aware of this (Samarin 1967b:108–112). According to Harris and Voegelin (1953:67), covert elicitation is beneficial: "Upon proper informant-linguist situations, we can trust the informant to give us disinterested responses because the informant is unaware of our strategy."

There is an ethical problem here. At worst, the fieldworker can be perceived as deceptive, or at best, hurting the consultant's feelings. On the other hand, all elicitation is covert to the extent that it is impractical and counterproductive for the fieldworker to tell the consultant every detail of what s/he is doing and looking for at every moment. And of course, the fieldworker must often write things down during the course of a fieldwork session that might not – in the consultant's opinion – be part of the linguistic information provided. Some consultants like to look over the fieldworker's shoulder to see what s/he is doing, or to make sure that s/he is not writing something different than the information s/he has provided. This is a case where working with illiterate consultants can be an advantage. There is no hard and fast solution to the ethical problem. There must be some basic trust between the two parties, as well as a basic understanding that in grammar gathering, lack of awareness does not entail deception.

Assuming there is no perception of deception, covert elicitation can be used when the speaker is bored by a repetitive or technical elicitation session, or when the speaker has strong feelings about what is important and interesting in their language. S/he therefore might resist certain types of elicitation. (This is more of a problem in syntactic elicitation than in lexical and morphological elicitation because most consultants are interested in words – a salient object cross-linguistically – and are more likely to be happy to provide them).

In such a situation, the fieldworker might try to build the sentences for elicitation around linguistic matters that generally interest or amuse the consultant. These could be things such as baby-talk or motherese, ideophones (Samarin 1967a, 1970, 1974, 1991; Hinton et al. 1994), personal names or place names, proverbs (Bryant 1945), riddles (Yvonne Treis, p.c.)[10], metaphors (Rupp 1974), figures of speech or idioms (Leaders 1991, Wilson 1990), language games (Bowern 2008:82), play languages (Kirshenblatt-Gimblett 1976, Sherzer 1982, 2002), slang (and, with extreme caution, swearwords (Vaux et al. 2007:292–295)), or various linguistic ideologies and prejudices. Or one can use Paradigmatic substitution elicitation (Section 12.2.2.3), in which case the fieldworker is less interested in the blank, and more interested in what is around the blank. The fieldworker should also be interested in baby-talk, ideophones, personal names, etc., in addition to what is in the blank. It is more a question of distracting the consultant so that relatively natural data are obtained with minimum boredom.

Paul (1953) mentions anthropological fieldwork done in Nigeria where the fieldworker instigated strong arguments between consultants in order to get more information. This bullying technique could be considered a form of covert elicitation. This is only worthwhile if the fieldworker knows his/her consultants and their

[10] Proverbs and riddles have the advantage of being short texts, and so are also useful for syntactic study, if the fieldworker looks out for any unusual syntactic or pragmatic features they might show.

culture extremely well. However, by definition, the fieldworker does not know the culture as well as the consultants do. So the bullying technique is mentioned here for the sake of completeness, rather than as a serious elicitation method.

12.2.2.10 Meta-Elicitation (Difficult to Perilous)

It is true that during early stages of elicitation, the fieldworker should not expect the consultant to provide analyses. If the consultant does offer an analysis, it should be noted down and used as a clue in the fieldworker's own analysis. But at later stages, once the fieldworker is sure of a particular consultant's linguistic sophistication, s/he can ask for simple types of analysis. Hale (1965:115), for example, talks about referring to a text and asking the consultant to analyze complex sentences. For instance, the fieldworker might ask the consultant what the two clauses are that make up the sentence. Or it might be reasonable to ask the consultant directly if there is a semantic difference between two constructions. Usually the consultant cannot tell the fieldworker, but it is worth seeing if s/he has a take on it.

It is important for the fieldworker to keep analytical questions simple, and not base them too heavily on theoretical concerns. The more complicated the question, the more likely the consultant is to misunderstand it. The fieldworker might know that Binding Conditions have been shown to hold in many languages, and if the consultant is very linguistically sophisticated, s/he might know this too. What needs to be avoided is for the fieldworker to look shocked if the consultant's analysis implies that Binding Conditions do not hold in his/her language (Vaux et al. 2007:264–266). Finally, the fieldworker not only has to know the consultant well, but also has to know the language well enough to evaluate the answer s/he is getting (Kibrik 1977:58).

12.3 Morphology Gathering Techniques

Two introductory remarks are in order. First, we repeat here the importance of looking at elicitation guidelines based on previous work on other languages of the area or in the same family. For example, reviewing the morphological marking possibilities for Australian languages listed in Sutton and Walsh (1979:25–31), or the possibilities listed for South Asian languages in Abbi (2001:109–112) would save the beginning Australianist or South Asianist considerable time.

Second, there is a tendency to think that it is possible to understand the syntax of a language without studying its morphology. This is only true for those rare languages where there is no morphology to speak of. In all other cases, the fieldworker's understanding of the syntax of the language is not going to be worth much until s/he has got a good understanding of the morphology. But it is also true that the fieldworker's understanding of the morphology is not worth much until s/he has a good understanding of the syntax.

So the fact that our discussion of methodologies for morphological elicitation comes before methodologies for syntactic elicitation (Section 12.4.2) does not mean that one has to be carried out and completed before the other. It is just for the sake of exposition that we list morphological elicitation techniques separately and before syntactic elicitation in this chapter.

12.3.1 Schedule-Controlled Elicitation for Morphology

Paradigm elicitation may be the most stereotypical activity of the fieldworker interested in morphology. However, in some traditions paradigm elicitation is not an obvious activity. So a Chinese linguist might omit the elicitation of paradigms simply because the concept of paradigms is not part of Chinese traditional grammar (Randy LaPolla p.c.).

But in general it is understood that if a language has morphology, it has got to have paradigms somewhere. The stereotypical schedule-controlled elicitation method for morphology is the paradigm schedule, i.e. a schedule or template based on the patterns that we expect the language to exhibit.

If the target language has no subject or object agreement indicated on the verb, or if tense and aspect are indicated through adverbs or derivational morphology, then scheduled paradigm elicitation for agreement, tense, and aspect might well produce false paradigms. Sanskrit- or Latin-based grammars of the Tibeto-Burman languages of Northeast India list false case declensions, since the morphology in these languages essentially marks semantic role, rather than grammatical relations.

As another example, it will take analysis of some textual data for the fieldworker to recognize an eye-witness evidential affix which, in early paradigm elicitation, masqueraded as a past tense marker (Kibrik 1977:66).

Every fieldworker describing a language with morphology constructs paradigm schedules. S/he will need to continuously refine and correct these schedules before finally moving on to schedule-free elicitation. Thus the first paradigms elicited are trials that will need to be changed as more is learned about the morphology of the language. If the fieldworker has no idea about the variables that will organize themselves into a paradigm, s/he will not get a paradigm right and clean on one notebook page. Several tabular schemes will have to be tried out. There are limits, of course, as to how many organizational schemes can be represented on a single notebook page.

Scheduled elicitation of noun (or adjective) paradigms is generally relatively non-problematic, the main categories being factors such as number, possession, gender or noun class, and case. It should not take long for the fieldworker to figure out which of these categories are relevant. Of course, one should always expect the unexpected. Some languages have unbelievably irregular noun paradigms, to the point that one must ask for, say, the plural of every single noun elicited. Other languages, like Western Apache, do not seem have any plural marking at all, until one discovers one day that half a dozen nouns, generally referring to humans, do have plurals – and of course, fairly irregular ones.

Scheduled elicitation of verb paradigms is much more difficult to do efficiently, since the variables or dimensions of a verb paradigm are much harder to predict. Therefore it is necessary to move on to schedule-free elicitation of verb paradigms at a fairly early point.

The fieldworker must have a basic set of intransitive and transitive verbs from different semantic classes represented in his or her verb paradigm schedules. Using these verbs in simple sentences, s/he should vary values for TAM (tense, aspect, and modality) and for person, number, and gender for subjects and objects (Vaux et al. 2007:232–236). Foley (2002:134) suggests that both positive and negated sentence equivalents be elicited.

Particularly with verb paradigms, some consultants will catch on quickly to what the fieldworker is looking for; but many never will, and will tend not to stay within the paradigm. They will come up with unexpected and unrequested TAM or person forms, or will switch verb stems on the fieldworker. Some speakers will not give the fieldworker forms of the same word for very long, but will happily stray across the whole range of vocabulary, and will reluctantly be reined back in to the "boring" word under discussion.

Try as s/he might, the fieldworker will not be able keep close control over the data. This is the part about paradigm elicitation that is confusing for the consultant and taxes the patience of the fieldworker. Three pieces of advice are offered here.

First, the fieldworker should keep his/her cool no matter where the consultant leads.

Second, the fieldworker should leave a lot of empty lines between forms in his/her notebook, and should be ready to connect one form with the other with arrows. If there are too many arrows, it is helpful to use different colored pens. Paradigm elicitation in a notebook always turns out to be quite messy and unsightly, and we have always been suspicious (or maybe just jealous!) of fieldworker colleagues who quickly manage to put neat and clean paradigms into their notebooks. A useful discussion of paradigm elicitation on the notebook page is in Hale (2001:92–94).

Third, it is helpful to try to place information in predictable areas on the page so it can be accessed and transferred easily to a database program, and then sorted into various paradigmatic relationships. One can create one's own template with a program like TOOLBOX or FLEx (see Section 8.7), entering one verb per page and then sorting by various features, and continuing further elicitation on that basis. When paradigms are complex, it is advisable to experiment with organizing forms according to different fields until one is more certain of what the relevant fields will be. Then, when the fieldworker is ready for paradigm-filling elicitation (Section 12.3.2 below), the transfer of the paradigm to TOOLBOX can be carried out.

It is also a good idea to combine schedule-controlled paradigm elicitation with ancillary elicitation (Section 12.2.2.8). For example, the fieldworker could record paradigms for each verb that arises during translation of texts. In addition to providing useful context, combining text translation and paradigms will guard against consultant fatigue and boredom.

When consultants are bored with paradigms, funny things can happen. The funniest example of paradigm elicitation we have ever read is in Reynolds (1993), taken from Barnum (1893), about a missionary trying to get a paradigm from a bored Yup'ik Eskimo speaker.

12.3.2 Analysis-Controlled Elicitation for Morphology

In this Section we discuss the methodology of Paradigm-filling elicitation (Section 12.3.2.1), and discuss strategies for eliciting Bound and free morphemes (Section 12.3.2.2), Stems and roots (Section 12.3.2.3), Noun and adjective morphology (Section 12.3.2.4), and Verb morphology (Section 12.3.2.5). No attempt is made to classify these approaches as rigorously as in Section 12.2.2, or to grade them by difficulty, but clearly these methods are variants or expansions of analysis-controlled elicitation.

12.3.2.1 Paradigm-Filling Elicitation

Paradigm-filling elicitation is distinct from, and should not be confused with, schedule-controlled paradigm elicitation (Section 12.3.1). In schedule-controlled paradigm elicitation, the fieldworker is busy refining the parameters and dimensions of the paradigm. In paradigm-filling elicitation, the fieldworker already has a good idea of what the paradigm should look like, but does not have all the forms to complete it. Examples and advice concerning paradigm-filling elicitation are in Mithun (2001:38–40) and Bowern (2008:93–96). For personal and demonstrative pronoun paradigms, a good source is Bliss and Ritter (2001).

Paradigm-filling elicitation can easily go astray because the native speaker really must have a knack for thinking about word forms out of context and for picking out and analyzing pieces of words. Some consultants develop a feel for what the fieldworker desires in a paradigm: consistency, focus, the right amount of imagination, and varying only one thing at a time. For others, this is too much abstract thinking about a language. Misunderstandings abound.

I [de Reuse] have had speakers give me the singular, dual, and plural forms of a verb when I was asking for first, second and third person forms. Or, a few forms into the verb paradigm, the consultant would grumble, "Why do you keep asking me for the same word over and over again?". Dimmendaal (2001:69) had similar experiences. Another consultant politely complied with my verb paradigms, for days and days, and then a year later asked me pointedly, "Are we still doing that verb stuff?".[11] This was a clear hint that this consultant had had enough. In our experience, only slightly masochistic consultants who are linguists themselves enjoy paradigm

[11] Actually, a less proper word than "stuff" was used.

filling. Because paradigm-filling elicitation, like schedule-controlled paradigm elicitation, is tedious, and can be extremely taxing, it should be blended in with more enjoyable and relaxing tasks, such as word elicitation or text translation (Sutton and Walsh 1979:23, Dimmendaal 2001:69).

Despite these problems, paradigm-filling elicitation is sometimes necessary for languages that have regular conjugations or rich agreement morphology. One cannot randomly wait for the paradigm to be filled (Kibrik 1977:57). Some linguists have found eliciting paradigms to be an enjoyable and productive task, as George Van Driem (p.c.) in his study of Himalayan languages.

In the Indo-European tradition of inflectional grammar, we expect speakers to have some idea of how big a paradigm is, and we think that if they know the language, they will know every form in the paradigm. For instance, German or Spanish schoolchildren are expected, after some practice, to be able to recite a complete paradigm of a common verb. However, many languages have such extensive paradigms, and some forms are so rarely used that most speakers simply do not know what all the forms are. With polysynthetic languages, we recommend doing partial paradigms of a great number of items, and coming back to them later rather than trying to do a relatively full paradigm for each item at one sitting; that is simply too taxing on both consultant and fieldworker. With the verb paradigms of Eskimo languages – which are polysynthetic – this approach is necessary. When cornered (literally) in a little-known box of a paradigm, the consultant will say s/he is not sure, or will feel pressured to make up an analogical form to satisfy the fieldworker. In our experience, if there is a genuine paradigm, the forms do exist, but one might have to ask several consultants with good memories, or unusually fluent ones, to get at the very rarely used forms.

Ideally, every effort should be made to complete paradigms; but if it takes the fieldworker and speaker fifteen minutes or more to get just one extra form of the paradigm, efforts are probably better spent elsewhere. Of course, the fieldworker will try to fill in paradigm blanks with guesses, and then check these hypothetical forms by eliciting speaker judgments on sentences created using them. In doing so, one must be wary of what Healey (1964:15) calls "false forms" brought about by the speaker trying to fill in the blanks for the fieldworker. This dressage effect should be expected. Indeed, the most dangerous problem with paradigms was pointed out to us by Michael Silverstein (p.c.): "If you do paradigms, you get into paradigm regularization." Healey suggests watching for facial expressions or hesitation that might indicate that the speaker is making something up in order to be "helpful". To avoid paradigm regularization, Sutton and Walsh (1979) also propose checking paradigms with at least two different consultants.

12.3.2.2 Eliciting Bound and Free Morphemes

Languages vary broadly regarding how hard it is to distinguish bound and free morphemes. In some languages, it is easy to recognize underlying potentially free morphemes, and one can conclude that the rest are all bound morphemes. In other languages, it is the bound morphemes that jump out at the fieldworker, and one can then conclude that whatever is not recognizable as bound must be a free or potentially free morpheme. The methods of Paradigmatic substitution elicitation and Fieldworker-driven substitution elicitation (Section 12.2.2.3) should be used to investigate the meaning and distribution of bound morphology. A good survey of useful techniques for determining the meaning of morphemes (as well as of larger constructions) is in Matthewson (2004).

In many languages, as in Pomo (Hokan of California) (Sally McLendon, p.c.), speakers will not recognize a stem if its prefix is cut off. On the other hand, for some languages, speakers who are literate have a tendency to write productive prefixes as separate words. We have noticed this tendency in Bantu, Yucatec Maya, Nahuatl, and Western Apache.

Some types of bound morphology, even though exotic by Indo-Europeanist standards, are not too hard to discover. Morpheme-internal change – as in Afro-Asiatic languages, Nilotic, Penutian, Athabascan and English *foot/feet, mouse/mice* – is easier to discover than are complex fusion and irregular morphophonemics in affixation. But when the morpheme-internal change depends on subtle differences in vowel quality, quantity, or tone, it is easy to miss it altogether.

Suppletion and total fusion are always tricky, because they can be identified only after the rest of the morphology is fairly well understood. So it is always good to check and recheck postulated suppletion and total fusion. What looks like suppletion might be poorly understood morpheme-internal change, and what looks like fusion might be an unusual but regular type of morphophonological alternation.

Identifying morphemes in polysynthetic languages is not necessarily harder. In Eskimo the first syllables of a word will always be the stem, and everything else will be suffixes. But identifying morphemes is time-consuming and challenging in a polysynthetic language with prefixation and suffixation, especially when complex morphophonemics or fusion blur the boundaries between morphemes (as in the Algonquian and Caddoan families of North America).

In eliciting compounds, care needs to be taken to distinguish phonologically between a compound word and a phrase composed of two words. If a recognizable word is longer that what the fieldworker has established for stems, s/he should suspect that this is a compound. Needless to say, this rule of thumb does not work for polysynthetic languages, where words can range anywhere from one to more than twenty syllables long.

The meaning of a compound word is often easy to determine. But languages vary widely in terms of how easy or how difficult it is to determine the meaning of the parts of a compound. In Germanic languages, there is not much of a problem, but in many South East Asian and East Asian languages, it is difficult to identify the meaning of the parts because of rampant homonymy. An example is given in Section 12.2.1.

If the language has a literary or grammatical tradition, the consultant might be able to provide morpheme-to-morpheme glosses of words. This is, of course, a welcome help, but it is good to remember that it might not be any more reliable than the help that an educated (but non-linguist) speaker of English could provide to a learner of English.

In particular, the names that an educated consultant gives to bound morphemes might be derived from prescriptive textbooks often based on Latin, Arabic, Sanskrit, or colonial European languages. Those terms might be completely inappropriate for the language the fieldworker is studying. For example, a tense or aspect label like "past perfect" might not be a valid category for representing the formal distribution of a specific set of morphemes. Also, the fieldworker can prevent the description from being Eurocentric by not trying to teach his/her consultant this terminology. Fieldworkers working in the generative tradition, such as Hale (1965) and Wong (1975), recommend teaching the consultant the technical terms. It is true that some interested, college–aged speakers might benefit from this, but in our experience, the "eyes glazing over" phenomenon is more common. In any event, in naming bound morphemes, our main recommendation is to base analysis on their form and not their function.

Morpheme order can sometimes be figured out by studying the longest sequences of morphemes found in the fieldworker's database. However, shorter and intermediate sequences should also be studied, in order to find out how they are assembled, and what the immediate constituents are. A good method for determining morpheme order is Grimes' (1967) Position class analysis. More modern methods of studying morpheme order rely heavily on the results provided by computer applications such as TOOLBOX, but it is important to remember that such tools will not provide one with immediate constituent analyses of morpheme sequences.

We can offer several other pieces of advice on recognizing bound morphemes. Strange cases of morphological homonymy should be expected. We all know the case of the English morpheme that has a clear allomorphic variation: /-z/ ~ /-s/ ~ /-əz/. As is well known, /-əz/ is the allomorph found after sibilants and shibilants, /-s/ is the allomorph found after voiceless obstruents (other than sibilants and shibilants), and /-z/ is the allomorph found elsewhere. "Surely", a linguist from Mars will say, "there cannot be more than one morpheme with these unusual allomorphs!" But as any English speaker knows, there are no less than three homonymous morphemes with this set of allomorphs: the plural, the Saxon Genitive, and the third person singular present indicative of a verb! Yet we instinctively avoid believing that this sort of homonymy could occur in an otherwise wonderfully complex language of the Amazon or Papua New Guinea.

Long inflectional affixes are easy to overanalyze, particularly in agglutinative languages. For example, the Southern Peruvian Quechua – *manta* 'ablative' looks like the dative – *man*, followed by the accusative – *ta*, and this might well be the proper analysis diachronically. In Manipuri (Chelliah 1997:250), the ablative – *təgi* looks like the locative – *tə*, followed by the genitive – *gi*. At first sight, little is gained by adopting such analyses in a synchronic description, but it is always a good idea to mention interesting diachronic facts in a descriptive grammar.

Synchronic double case marking and case stacking (or Suffixaufnahme) exist in some languages, as discussed by Plank (1995).

Sometimes a fieldworker might consider two elements to belong to two mysterious morphemes, until s/he realizes that they are allomorphs of the same morpheme. For a long time, I [de Reuse] was puzzled by the element /-ftə-/ and by the element /-pətə-/

in Siberian Yupik Eskimo verb morphology. Speakers could not explain what either meant at all. After a few months, I realized that they were allomorphs of a common morpheme: /-ftə-/ occurs after vowels, and /-pətə-/ occurs after consonants. An alternation /-f-/ ~ /-pə-/ is not common in Eskimo morphology, but it is not unexpected, either. Having now twice the amount of morphological material to look at, I figured out that this was an inferential evidential meaning ‘evidently...’ or ‘I infer’ (de Reuse 1994:169).[12]

What to do when two formally identical morphemes occur in the same word? If they clearly mean two different things, there is no problem; this is a case of two homonymous morphemes following each other. But if the fieldworker is not sure of the meaning of the morphemes, there are several possibilities. It could be that the same morpheme occurs twice. In this case, they could be in sequence, or might sometimes be separated by other material. If they always occur in sequence, it could be (a) repetition for emphasis, as in *very very good*; or (b) reduplication, which might have several specific meanings. If they are not always adjacent to one another, it could be that the same morpheme is repeated recursively, with scope differences. Such cases can be found in Siberian Yupik Eskimo. In (2), the element -*sqe*- ‘ask to V’ is used recursively, with scope differences.

(2) iitghesqesaghiisqaa

itegh-	-sqe-	-yaghtugh-	-sqe-	-aa
come.in	*ask.to*.V	*go.to*.V	*ask.to*.V IND.	$3s_i.3s_j$

‘He$_i$ asked him$_j$ to go ask him$_k$ to come in’

12.3.2.3 Eliciting Stems and Roots

Eliciting stems and roots in isolating languages is easy. Regarding morphologically complex languages, the fieldworker can take the following approach to gathering noun and verb stems or roots. Once the basic stem or root phonotactics are understood, it is possible to elicit other stems or roots by constructing potential nonce forms based on the stem’s phonological structure (Crowley 2007:110–111) or based on its rhymes (Dimmendaal 2001:66–68). One might inflect them minimally, and ask the consultant to provide new stems or roots. This is a long and arduous task. The idea is to jog the consultant’s memory. S/he might respond with the form fieldworker expects to exist, but not very often. The consultant might say: “no, but there is this one that sounds like it”, and if the fieldworker is lucky, that something else is a form that is new to him/her. This approach was fruitfully carried out by Laughlin (1975) for the Mayan language Tzotzil, by Kaufman and Justeson (2003:6–7) for Mayan and other Mesoamerican languages, and by Krauss (2006) for Eyak of Alaska.

[12] It is also interesting that one speaker combined the two allomorphs in post-consonantal position (as -*pətəftə*-), but retained -*ftə*- after vowels (Badten et al. 2008:609).

In languages with a rich morphology, it is also often the case that verb stems or roots are harder to elicit than noun stems or roots, because verb morphology is generally more complicated than noun morphology. In phonological or phonetic fieldwork, it is easier to work with minimal sets of stems or roots; thus it will be more difficult to work with verbs than with nouns (Maddieson 2001), and more preparatory morphological analysis will be required. Some languages have citation forms which are identical to or close to a stem or root. In other languages, the citation form might be provided in the infinitive, but the infinitive might be quite a bit more complex than the stem or root it contains.

It is generally easy to assign some meaning to a noun or verb stem. But in Athabascan languages, the stem meaning can be so general that a better understanding of the morphology is gained by noting that certain affixes (called "thematic affixes" in Athabascanist terminology) are as much a determinant of the whole meaning of the word as the stem itself. This is fortunately not a common situation. More often, the principal challenge is to identify and define each bound morpheme, as discussed in the previous section.

12.3.2.4 Eliciting Noun and Adjective Morphology

As mentioned in Section 12.3.1, the most common and regular categories to check for are gender or noun class, number, and case. Working with fusional languages is tricky because more than one nominal inflectional category can be subsumed under one morpheme. A nice example from Polish can be found in Vaux et al. (2007):215).

For gender or noun class morphology, different nouns can be collocated with different modifiers; e.g. "tall boy, tall girl, thick book, long pencil", and so on. One can also collocate adjectives with names – one male name, and one female name.

For number morphology, the situation is either easy or tricky. If there is obligatory plural or dual inflection, it will show up quickly in the data, and will be quickly identified by both consultant and fieldworker. If the number marker is rare or optional, elicitation cannot be relied upon, so number marking will have to be studied in texts. Hän Athabascan has a rather optional number marker, which had always been translated as plural, until one day it occurred in a text added to "my father". Then it dawned on me [de Reuse] that it was not a plural marker at all, but that it means something like "and company", "and group", "and them". Examples of suspicious "plural" markers that turned out to be collectives or distributives are discussed in Mithun (2001:40–42).

For case morphology, one should collect nouns in various semantic role positions – agent, patient, etc. – and use them with verbs with different valency values. The fieldworker should be aware of morphological elements that look like or have been explained as case marking. One should begin with the idea that case marking is obligatory, exceptionless, and paradigmatic. Anything that does not have these attributes is just not case marking, or it is case marking overlaid by some other system, such as a pragmatic system (Barðdal and Chelliah 2009). To unravel a system like case morphology, a lot of Schedule-controlled paradigm elicitation (Section 12.3.1) and Paradigm-filling elicitation (Section 12.3.2.1) might be necessary. For locational case morphology, physical objects can be placed in view. Then the fieldworker can indicate locations above, below, and to the side of these objects. (See the discussion of Prop-driven elicitation in Section 12.2.2.2). Finally, the fieldworker can refer to objects which are not in view, but are manifest in the material culture.

Elicitation of nominal morphology is further discussed in Vaux et al. (2007):216–217).

Adjective morphology (assuming the language does have adjectives) can be studied by collocating adjectives with different nouns to test for agreement. In adjectives, the basic morphology will involve equatives, comparatives, and superlatives. Such forms can be elicited by drawing lines of different lengths, or showing the consultant strings of different lengths.

12.3.2.5 Eliciting Verb Morphology

We begin with a discussion of Tense-Aspect-Mood (TAM) elicitation, which is one of the most difficult kinds of morphological elicitation there is. The fieldworker should not try to understand TAM only through translation, but should always use textual information to check hypotheses. S/he should be careful, of course, about setting up paradigms using Indo-European categories for TAM. These may over- or under-differentiate the distinctions made in the target language.

Indo-European languages – particularly Slavic languages – typically have Indicative, Subjunctive, and Imperative moods, and Imperfect and Perfect aspectual distinctions. As far as tense is concerned, modern Indo-European languages typically mark several pasts, often one periphrastic present, and morphological and/or periphrastic futures.

One might think that the problems of TAM elicitation are less prominent if the native speaker is aware of a grammatical tradition, or has experience with a written language. We have found that such background helps very little; the consultant might be aware of an Indo-European or other non-indigenous grammatical tradition, and as a result might also be guilty of over- or under-differentiating the distinctions made in the target language.

Furthermore, there are plenty of TAM quirks that are not clearly part of any grammatical tradition. Consider the use of the present tense form for future reference in English, as in *She goes to London tomorrow*. Or the past tense form for future reference in Dutch: *Je kwam toch morgen?*, literally something like “You came tomorrow, right?”, which actually means ‘You said you were going to come tomorrow, right?’.

Even linguistically sophisticated speakers, who have a good sense of formal devices and might even be able to label them, will have trouble explaining how and when TAM marking is used. I [de Reuse] have known all my life that my native language, Dutch, has two past tenses. One is called the Onvoltooid Verleden Tijd (OVT), or 'noncompleted past'; often translated as Imperfect, this is the original past, similar in form to the simple past in English. The other one is called the Voltooid Verleden Tijd (VVT), or 'completed past'; often translated as Perfect, this is similar in form to the English Present Perfect, though it does not function at all in the same way. To this day I do not know why I sometimes use the OVT, and sometimes the VVT, and the textbooks are not clear about the distinction either.

Under these circumstances, one can hardly expect a linguistically unsophisticated consultant to be able to explain the difference between two TAM forms. For the consultant to translate TAM from a contact language, s/he would have to know the nuances of the contact language, avoid literal translations, and be able to provide the exact TAM equivalent in the target language. The fieldworker is encouraged to try doing this himself/herself for two languages that s/he knows, and then s/he will realize how difficult the task is.

The difficulty is compounded when TAM morphology is organized in paradigms where the TAM distinctions are intertwined or fused. And that is very often the case.

Evidentiality, regardless of whether it is morphological or marked in another way, is particularly low on a "hierarchy of elicitability" (Silverstein 1979:234). Even though evidentiality is a category logically distinct from TAM (Aikhenvald 2004), one of the main reasons why it is often so hard to recognize and elicit is because it is so often intertwined or fused with TAM morphology (Mithun 2001:45–48).

As a result, while the translation prompt might entail manipulating one or two TAM features in the contact language, several other TAM (or evidentiality, or mirativity) features might occur in the equivalent target language sentence.

For example, in the target language, a third person pronoun subject might require the verb to occur with an evidential, or hedge (such as, "he says" or "evidently"), since in some languages speakers are less comfortable speaking about third persons.

As another example, in the attempt to translate a Past or Future from English into Western Apache, consultants often add the enclitics *ni'* or *doleeł*, respectively, after the verb. The fieldworker should not be misled into believing that *ni'* just means 'Past tense', and that *doleeł* just means 'Future tense'. In reality, these enclitics mark not only tense but also an epistemic modality (or validationality) of certainty. So *ni'* means 'past tense + certainty of speaker' and *doleeł* means 'future tense + certainty of speaker'.

A strategy often used to force different tense readings is to construct sentences with the adverbs "yesterday", "today", and "tomorrow" (see syntagmatic elicitation under Section 12.2.2.3). For a first pass at elicitation, this is useful. The more abstract or semantically complex the adverbs get, the harder the consultant's task becomes. Thus trying to elicit aspectual differences using adverbs such as "already", "just", or "only" is not advisable; the meaning the fieldworker is thinking of might not be the same as what the consultant is thinking of, e.g., *just* for 'right now' or *just* as in 'only VERB'. Also, these adverbs may elicit semantically complex concepts. For example, "already" could be translated as a perfect but might also include a manner component, i.e. refer to an action completed before it should have been completed.

Inflectional plural marking on verbs is usually a problem as well. Eliciting non-singular subject inflection is always harder than eliciting singular subject inflection, since a distinction has to be made in many languages between inclusive and exclusive "we", and between plurals and duals. Reverse translation of English "you and I" might yield a literal translation of "you and I", rather than the dual inclusive subject inflection that the fieldworker was hoping for. Reverse translation from English is a particularly acute problem with the ambiguously singular or plural *you*. If one asks for translations of "you all", "you bunch", "all of you", or "the two of you", one might get literal translations, rather than the "you (pl.)" or the "you (dual)" inflection hoped for. "They" tends to be easier, assuming the language has third person plural inflection, although here one has to look out for special distributive or collective plural forms.

12.3.3 An Illustration of Morphological Elicitation from Athabascan

Here we discuss some additional elicitation problems arising in a morphologically complex TAM and person system – that of the Athabascan languages – based on my [de Reuse's] fieldwork on two languages of that family: Western Apache of Arizona, and Hän of Alaska and the Yukon Territory. More on the joys and perils of eliciting Athabascan verb morphology are in Rice (2001:236–238).

First, some background. There are, simplifying somewhat, five basic TAM distinctions in these two languages: Imperfective, Perfective, Future, Optative, and Customary. Note that Imperfective, Perfective, and Customary are aspectual distinctions, Future is a tense distinction, and Optative is a mood distinction. In Athabascan these five distinctions contrast in a paradigmatic way, and to further confuse non-Athabascanists, these are all called "mode" distinctions in Athabascanist terminology. This is typical of Athabascan languages, but certainly not typical of languages in general. Thankfully, linguists preceding me had figured out these distinctions.[13]

Experience taught me that the Perfective is fairly easy to elicit by asking for the translation of "VERBed in the past", "it happened", "it is over with, etc". The consultant will be happy to volunteer a first or third person subject form, but not a second person subject form. We return to subject person forms later.

For the Imperfective, a prompt such as the English progressive "I am VERBing" generally works. But there are various problems here. Not all English verbs can

[13] One point that has not been made before is that the semantic contrast between Athabascan Imperfective and Perfective seems to me [de Reuse] to be very similar to the contrast between Imperfective and Perfective in Arabic. Now, that is my impression from my reading, but until an Arabist has a serious look at Athabascan aspect, we will not know for sure.

easily be put into the Progressive. So, *"I am liking it" will not work, but "I am eating it" will. So one is limited to verbs that go with the English Progressive. I also noticed that for the Imperfective, the consultant will be happy to volunteer a first person form, will offer a third person subject form with some difficulty, but will not readily provide a second person subject form. Again, more on this later.

In any case, it is clear that the Apache or Hän Imperfective and the English Present Progressive are not correlated one-to-one, and one should not expect them to be. Accordingly, there is a discrepancy between a translation from English to Apache or Hän, and a translation from Apache or Hän into English. When I ask for the English translation of an Apache or Hän Imperfective, I often get an English immediate future "I am going to VERB". However, beginning with an English immediate future prompt "I am going to VERB", the consultant will always give a Future form in Apache or Hän. Such a lack of correlation in target language translation and reverse translation is quite typical of TAM elicitation work.

As a result, for Apache and Hän at least, the Future is the easiest to elicit, since asking for a translation of an English Future will yield a Future in these languages.

The Optative is more difficult. English prompts such as "I wish to VERB" or "I want to VERB" result in literal translations that do not always contain Optative forms at all. For some consultants, translations from English of "I would VERB", or "I should VERB", or "Why don't I VERB?" will yield Optative forms. The most convenient way, though, for both Apache and Hän, is to identify a prohibitive particle that automatically selects a second person Optative verb form. So one can ask for a translation for "don't VERB". This works with most consultants, but some particles that mean "don't" select an Imperfective form, so the fieldworker has to be sure to find the particle that selects the Optative.

The Customary is easier too, because one can ask for a translation of "I VERB every day". or "I VERB every morning, or every week, etc." Of course, this will also give one translations of "every day", "every morning" etc., but if the language has a distinct Customary, this is where it will show up. The system is not foolproof, though, since some consultants will provide an Imperfective together with translations of "every day", "every morning" etc.

Thus, as illustrated by the situation in these two Athabascan languages, eliciting a TAM inflectional system is not easy. The problems depend on the specific language, of course, but the above paragraphs are representative of the pitfalls. There is a trial period of a few days before one discovers what works best. I spent about a week getting a lot of Perfectives and Futures from translations of English Pasts and Futures, and I got a variety of verb forms on the basis of translations from English Presents. Maybe even more daunting is the fact that English prompts that yield the right inflectional forms from one consultant might not work with another consultant. The process has to be re-started with every new consultant.

We now consider the elicitation of subject person morphology marked on the verb in these two Athabascan languages. With some consultants, getting subject person is very easy; one can ask for translations of "I am VERBing", "you are

VERBing", "s/he is VERBing", etc. But with others, as is well known from the literature (Basset 1951:19, Foley 2002:134, Bernard Comrie, p.c.) a shifting deixis problem gets in the way: "I am VERBing" translates as "You are VERBing", and conversely "You are VERBing" translates as "I am VERBing". And some speakers can do this for days, even after being made aware of the problem. If there is object pronoun inflection, the problem is compounded: "You understand me" will come out as "I understand you"; "I will give it to you" will come out as "you will give it to me", and so on.

As noted in my discussion of Apache and Hän above, TAM contrasts also influence the ease or difficulty of eliciting subject person morphology. As mentioned, Perfectives are easiest to elicit with first or third person subjects, Imperfectives and Futures are easiest to elicit with first person subject, and Optatives easiest with second person subject. We now consider why that would be the case. It is not that the paradigm does not have these forms; it is just that various pragmatic and cultural constraints are at work.

Perfectives are easiest to elicit with first or third persons, and difficult to elicit in second person, because Perfectives refer to past events, and an Apache or Hän speaker does not like to be in the position of telling an interlocutor what the interlocutor did. The speaker simply does not have the epistemic or evidential authority to make assertions about the actions of the person s/he is talking to.[14] Instead, they prefer talking about what they themselves did or what a person not present did.

Imperfectives and Futures are easiest to elicit in the first person because they refer to uncompleted actions or events, and again, the Apache or Hän speaker does not have the epistemic authority to predict that the interlocutor or someone else is going to do something. Since the Future tends to mark absolute certainty as well as tense, there are cultural constraints against using it in subject persons other than the first person.

As explained above, it is easiest to elicit Optatives in the second person subject. There is also, however, a fairly easy way of getting a second person subject Imperfective. For both Apache and Hän, it appears that a semantic imperative – i.e. a command in the second person – is a formal Imperfective, so if one asks for a translation of a command in the second person, one almost always gets a second person Imperfective. Perfectives in the second person are always a bit uncomfortable to elicit, and one has to set up a scenario where someone can remind someone that "You VERBed this!", without making it sound like an accusation!

Third person subject and third person object forms are also confusing for many consultants. Rather than ask for "s/he understands him", and the like, it is better to start with sentences with explicit and easily recognizable proper names, i.e. "Paul understands John", etc. For languages of India, Abbi (2001:141) suggests using kinship terms as subject nouns as a good way to see if there is honorific marking on the verb.

Needless to say, all of the above are relevant to Reverse translation elicitation (Section 12.2.2.6). In Target language translation elicitation (Section 12.2.2.4),

[14] Nancy Caplow (p.c.) observed the same reluctance during her fieldwork on Tibetan languages.

a second person subject Imperfective form will not even be recognized as Imperfective, but will be translated into English as a command. Some Apache speakers, for reasons that are not entirely clear, will translate a third person subject verb form into an English "–ing" form. If this is done consistently, the fieldworker should not dwell on it, but should catch on, and be thankful for a consistent clue.

12.4 Syntax Gathering Techniques

The fieldworker need not worry about narrow transcription for syntax, as long as everything is recorded so that s/he can do a narrow transcription later on if needed. However, the fieldworker needs to have figured out the basic phonemes and the basics of stress and tone. It is impossible to carry out a reliable syntactic analysis without this preliminary phonemicization, regardless of how tempting it is to skip the phonemicization business altogether and move on to the potentially more exciting business of syntactic elicitation. The fieldworker just cannot rely on consultants telling him/her that these two words/clauses/sentences sound the same or not, if no phonological analysis has been carried out.[15]

Wong (1975) is one of the rare early discussions of the relationship between Chomskyan grammar and fieldwork. He states that, for generativists, the goal of fieldwork is to discover the internal grammar of the speaker. Because the goal is to gain insights into an internal grammar, the main fieldwork methods to be used are informant judgments and reactions based on introspection (Wong 1975:48–50). The theory can influence the order in which information is collected: if the fieldworker considers syntax to be basic, s/he can do this first, and then phonology can follow (Wong 1975:51). In our opinion, Wong's view that phonemicization can be carried out after the basic syntax has been elicited is naive.

It is not uncommon for consultants to tell the fieldworker that "two words mean something different but sound the same", when in fact the words clearly do not sound the same. There might well be a tone difference which the consultant has chosen to ignore, because for him/her, "sounds the same" really means: "the vowels and consonants sound the same". The fieldworker needs to be aware of this phonetic issue for morphological and syntactic analysis.

It is also not uncommon for the consultant to be aware of a subtle phonemic distinction in texts, but when repeating the sentence or clause from the text in isolation, to not produce the sound that occurred in the original sentence. This could be because the sound only makes sense when the sentence is in its context. So even when working with texts, the fieldworker has to make sure that the material used is repeated exactly the same way.

---

[15] Once a preliminary phonemicization is carried out, it might well be possible for the syntactic fieldworker to omit, for some time, certain aspects of the phonology in transcription, such as perhaps tones, gemination, or vowel length. But in the final analysis and description, a full transcription should be presented.

Syntactic elicitation is different from morphological elicitation in that it there is so much to ask that it is easy to forget about relatively simple syntactic constructions. Yes/no questions, for example, are among the easiest things to elicit, but in Hoijer's extensive work on Chiricahua Apache (1938, 1946), as well as in Hoijer's unpublished notes, there is nothing about yes-no question marking.

Also, no matter what method of syntactic elicitation is used, there are lots of syntactic topics that are hard to elicit reliably. A partial list includes: conditions on Wh-movement; definiteness and specificity; parasitic gaps; possession of abstract nouns such as 'his honesty'; quantifiers (other than numerals); relativization; scope; and weak crossover. The fieldworker should also look out for clause-chaining. There is a temptation to analyze texts as containing short sentences, but this is not possible in languages with clause chaining (Himmelmann 2006:270–271).

Some consultants are good at syntactic tasks, while others do not have a clue, so it is necessary to test a potential consultant in advance. It is also good to learn to transcribe a good bit of text before attempting syntactic analysis (Sutton and Walsh 1979:32). It is important not to start with syntactic elicitation right at the beginning of a field session. It takes about an hour to settle into the consultant's true variety to get the most natural syntax from them. Before that, as Milroy (1987) notes, syntax is highly variant, and it is hard to pin a particular variety or syntactic system down before this time.

Guides to experiments in syntax, some of which the fieldworker can attempt, are Cowart (1997) and Crain and Thornton (1998). Useful exercises in syntactic elicitation are in Vaux et al. (2007):284–286).

12.4.1 Schedule-Controlled Elicitation for Syntax

We cannot think of any reliable schedule-controlled elicitation techniques designed specifically for syntax. Many syntax elicitation schedules simply ask the fieldworker questions about syntax, such as "what are the focusing strategies of the language?" (see Section 12.2.1), so the fieldworker has to resort to analysis-controlled elicitation.

Schedules that require one to reverse-translate large numbers of complex but syntactically interesting sentences into the target language have not been published, to our knowledge. This is a good thing, because reverse translation is quite unreliable for advanced syntactic research, as further discussed in Section 12.4.2.4.

12.4.2 Analysis-Controlled Elicitation for Syntax

The topics we will discuss in this section are Target language manipulation (Section 12.4.2.1), Target language translation (Section 12.4.2.2), Target language construction and introspective judgments (Section 12.4.2.3), Reverse translation

(Section 12.4.2.4), and Ancillary elicitation (Section 12.4.2.5). Again, no attempt is made to classify these methods as rigorously as in Section 12.2.2, or to grade them by difficulty, but clearly these methods are variants or expansions of analysis-controlled elicitation.

12.4.2.1 Target Language Manipulation in Syntax

Three elicitation techniques useful for syntax are described under this heading.

Word Movement Elicitation

This approach is of course useful for figuring out word or constituent order, but the fieldworker cannot entirely rely on it, because consultants often insist on fewer ordering possibilities than are found in texts. Word movement elicitation is also useful for testing the scope of a quantifier (such as English *all*, *some*, *none*, or *many*) (Vaux et al. 2007:133–134). The fieldworker produces sentences with scope in different places, and asks what the meaning is. In eliciting interrogation, one can use elicited question words to form interrogative sentences, or can put the question words at the beginning of the sentences, since that is a cross-linguistically common way of forming questions. Then the fieldworker can try moving the question-word further to the right (Abbi 2001:179–180).

Sentence Combination Elicitation

Hale (1965:114) suggests asking the consultant to combine two sentences. This can be useful in figuring out the structure of conjoined, subordinate, and relative clauses; clause chaining; ellipsis; and antecedent reference. To get conjoined or subordinate sentences, one can suggest two semantically related sentences and then ask the consultant to conjoin them (Abbi 2001:202). Examples are in (3) and (4):

(3)	Stimulus:	I went to the store. At the store I bought some candy.
	Response expected:	I went to the store and bought some candy.
(4)	Stimulus:	Ram wore his sandals. Ram did not wear his brother's sandals.
	Response expected:	Ram wore his sandals but not his brother's.

Noonan (1985:137–138) gives useful suggestions about how to obtain information on complement systems by creating sentences. However, it remains imperative to study complex sentences in texts (Foley 2002:136–137), as discussed in Section 12.4.2.5.

Intonational or Stress Elicitation

Although languages vary widely with respect to how intonation functions, it is always worthwhile to try taking a declarative sentence already elicited and to repeat it with rising intonation. Alternatively, taking an interrogative sentence with a rising intonation and repeating it with a falling intonation might also draw interesting feedback from one's consultants.

Varying the stress of a phrase can also lead to interesting results. Consider the following example from Dutch, in (5–8). These are written, in the regular spelling, except that in (7–8) an extra loud stress, louder than normal word stress, is marked with the acute accent:

(5a)	De hond beet de kat.	'The dog bit the cat.'
(5b)	De hond beet de kat.	'*The cat bit the dog.'
(6a)	De kat beet de hond.	'The cat bit the dog.'
(6b)	De kat beet de hond.	'*The dog bit the cat.'
(7a)	De hónd beet de kat.	'It is the dog that bit the cat.'
(7b)	De hónd beet de kat.	'It is the dog that was bitten by the cat.'
(8a)	De kát beet de hond.	'It is the cat that bit the dog.'
(8b)	De kát beet de hond.	'It is the cat that was bitten by the dog.'

Any sort of elicitation might yield (5a-b) and (6a-b), from which the fieldworker will correctly conclude that the word order of such simple Dutch sentences is Subject-Verb-Object, and that Object-Verb-Subject is not a possible order. However, if the fieldworker adds extra loud stress to the first noun, she or he might come across sentence pairs like (7) or (8), which are ambiguous in Dutch. In (7b) and (8b), the word order is totally unexpected. Hopefully, the fieldworker will realize that this extra stress (which we could call focus stress) has something to do with the unexpected meaning or word order.

Then the fieldworker can test his/her hypothesis by using the focus stress on other sentences, and come to the realization that sentences with focus stress are ambiguous in Dutch. The fieldworker might well have to construct contexts for the sentences with focus stress, and from my [de Reuse's] experience it might take a while for a linguistically unsophisticated speaker of Dutch to come to terms with this important fact about Dutch word order, focus, and stress, since it is not part of traditional wisdom on Dutch grammar.

12.4.2.2 Target Language Translation in Syntax

When asking the consultant to translate individual sentences from target language texts into English, one might notice strange things, even when the consultant has good control of the contact language. In Object-Verb languages like Western Apache and Hän, simple sentences are often translated in a very roundabout way. So the equivalent of: "the cat the dog bit" meaning 'the cat bit the dog', comes out as something like: *the cat, that is who the dog was bitten, by*. Similarly, "the man the lake to went", meaning 'the man went to the lake', comes out as something like: *the man, the lake, to, that is what he went, to*. But actually, there are no passives or pseudo-clefts (Payne 1997:279) in the original sentences. What happens is that the consultant is thinking about the sentences bit by bit and tries to keep the Object-Verb word order in the translation. This happens with consultants who know that they cannot translate literally, and so they know that *the man, the lake, to, he went* will not be sufficiently comprehensible. So translation from the target language to English gives the fieldworker only a rough idea of what the sentence contains.

In the study of constructions that a language does not have, it is sometimes helpful to translate into a third language, because this forces a longer explanation. For example, many languages do not have ideophones, but translation into a language which has no ideophones is more helpful than translating into a language with ideophones (Samarin 1967a). When both the target and contact language have ideophones, a clever speaker will translate the ideophones, finding a match from the target to the contact language but not providing much information on what the ideophone does or when it should be used. In these instances, it might be useful to use a different contact language, one that that does not have ideophones, to stimulate a longer, wordier, more circumlocutional and more informative translation. We are using ideophones as an example, but this idea has relevance for any syntactic construction where the target and contact languages are very close.

12.4.2.3 Target Language Construction and Introspective Judgments in Syntax

What was said in Section 12.2.2.5 applies here, too, with a vengeance: syntactic grammaticality judgments are generally less reliable than phonological or morphological judgments. Useful studies about judgments of syntactic acceptability are in Greenbaum (1973, 1976, and 1977), and an excellent overview of the studies of introspective judgments of syntactic acceptability is in Diller and Khanittanan (2002).

No doubt, all English speakers will agree that "Chomsky cat the killed" is not a possible English utterance, and all English speakers will agree that "Chomsky killed the cat" is a good utterance.[16] But these are not the sorts of things syntacticians look for in fieldwork. These simple problems are supposed to have been solved.

The surprising thing is that often very simple sentences will be rejected. Speakers may give judgments without thinking of all possible pragmatic situations where a construction could be used. Therefore, it is important to provide some real-world context (Vaux et al. 2007:253–258). An English example of a sentence

[16] Unless, perhaps, they really like cats, or Chomsky, or both!

rejected by some is “Birds fly.” Indeed, when would one ever feel the need to say something like that? Even a more natural sentence, such as “I see a dog” might well be rejected with the comment that no one says such things, since “we see dogs all the time here” (Anthony Woodbury, p.c.). In this case, the fieldworker might want to construct a real word context and then say something like: “I am looking for X’s dog, because I am supposed to go hunting with it. I cannot find it. Someone who doesn’t know which dog I am looking for might say: “I see a dog.” Then I might say, “Oh yes, but that is not the dog I am looking for.””

True, a syntactician might not ask for something as silly as “I see a dog” in the first place. It is the hard, diagnostic cases that s/he is interested in, things like parasitic gaps, “picture of him” sentences, and that sort of thing. There is, of course, some justification in constructing sentences with rare syntactic phenomena such as parasitic gaps, just to check them out (Bjerre et al. 2008:159–160). If these things exist, the speaker might well have intuitions about them. But many linguists have a hard time imagining how difficult it is even for a linguistically sophisticated native speaker to know what to do with such sentences. I [de Reuse] have been used by Dutch formalist linguists as a guinea pig for parasitic gap sentences in Dutch, since I am a native speaker of Dutch. I am linguistically sophisticated. Can I tell you whether some of the Dutch sentences are grammatical or not? I cannot process them. I just have no grammaticality judgments on some of those things.

Similarly, sentences like the ones given under (10) below might tell the syntactician something interesting, but they require superhuman skill for a consultant to process. See also (Vaux et al. 2007:250–258) on this problem.

The formal syntactician might respond that these are pragmatic problems. The fieldworker might respectfully agree with the speaker that this sentence is indeed pragmatically weird and that s/he agrees that no one would ever say that, and then s/he might ask the consultant, “Can you tell me anyway whether it would be grammatical if it was not pragmatically weird, please?” There must be better ways of doing syntactic analysis than having to use this sort of cajoling of the consultant.

The formal syntactician might also counter that the target language construction and introspective judgment method is everyone’s favorite for rapidly obtaining useful negative data, i.e. ungrammatical sentences, indicated in the literature with a preceding asterisk. However, it is not absolutely necessary, in empirical linguistic research such as corpus linguistics, to obtain negative data, as argued by Sampson (2002). Also, good descriptive fieldwork will result in plenty of negative data through the mistakes of the fieldworker, which the consultant will point out. If the fieldworker feels the need to gather large amounts of negative data, that is probably because the consultants do not react appropriately to the fieldworker’s mistakes. In this case, the fieldworker needs to work on a more productive relationship with the consultant. And if the fieldworker makes no mistakes, then something is really, really wrong!

Furthermore, consultants with no linguistic sophistication cannot expected to be familiar with the idea of right, wrong, and possible grammar. Assuming, as many syntacticians do, that there is such a thing as a cline of acceptability, the concept of such a cline will be foreign to consultants as well. So getting them to put OK, or an asterisk, or one or more question marks in front of a sentence is simply not a reliable procedure.

All these cautionary remarks are not intended to completely discourage the descriptive fieldworker from using the introspective judgment method. So we will now provide a few pointers about elicitation of grammaticality judgments in syntax**.**

To prevent "priming" one's consultants, one should mix grammatical with ungrammatical sentences. If the fieldworker provides a whole list of all expected grammatical or all expected ungrammatical sentences, the consultant will be stuck in a pattern of answering and might overgeneralize (Vaux et al. 2007:254, Bowern 2008:101–102).

Obviously, non-leading questions must be posed. The fieldworker should not say: "That one is good, right?" or "This one is bad, right?" The fieldworker should also learn to keep a poker face, and not show approval or disapproval of grammaticality judgments (Vaux et al. 2007:27–28, 255–256, Bowern 2008:90).

Letting consultants talk about their language, and about the grammar of their language, if they are inclined to do so, is of course a good idea. This way, consultants can become interested in things about the language. For example, the ordering of adjectives in an English noun phrase is often of interest to speakers.

12.4.2.4 Reverse Translation in Syntax

In reverse translation, the fieldworker offers a sentence in English (or other contact language), and asks the consultant to translate it into the target language. This method remains the most popular method for beginning syntacticians, but it is also the most fraught with difficulties and misinterpretations.

When using reverse translation, it is particularly important to realize that complex sentences with real but subtle differences, and sentences with repetitious full lexical items, are going to be quite confusing to the consultant. It would be hard for any consultant without at least a high school education to give reliable translations of the sequence in (9). A few of these can be asked at a time, then another task should be carried out, and of course comparisons with equivalent sentences in texts should be made.

(9) I know (that) you saw Mary.
I know whether (or not) you saw Mary.
I know (the one) who(m) you saw.
I don't know (that) you saw Mary.
I don't know whether (or not) you saw Mary.
I don't know (the one) who(m) you saw.

In some cases, reverse translation will not even work with constructions considered easy to elicit. Sandalo (1997:74) in her grammar of Kadiwéu, a Waikurúan language of Brazil, states: "I could not find any simple yes-no questions in Kadiwéu. When I tried to elicit questions such as "Did you find clay/Have you found clay?", they gave me sentences such as "When did you find clay?"" We have no doubt that Kadiwéu has a strategy for forming yes/no questions; it is just that this particular elicitation method failed to reveal them.

Although this should be obvious, the fieldworker should never try reverse translation of sentences that are hard to interpret even in the contact language. Examples of such sentences are in (10).

(10) There was a woman that Paul told Betty that the man kissed (her).
This is the man which I saw the boy that the girl who introduced him to her is snoring.
Whoever stares at which man's wife will be executed.
Which knife did John kill the cat with and Paul stole?

A consultant with a very kind disposition might be able to make something of such sentences, but most will just be irritated, and translations will likely be calques.

12.4.2.5 Ancillary Elicitation in Syntax

The use of texts as a means of avoiding the unreliability of grammaticality judgments in syntax is discussed in detail in Chelliah (2001:158–161).

Eliciting discourse-related syntactic matters such as topic and focus marking is also best done through texts. In recorded narratives and conversations, the fieldworker can change the order of constituents, or can change stress, length and pitch (Abbi 2001:178). The Dutch example with focus stress (5–8) discussed above is the sort of thing that could also have been discovered in a text first.

Particles (cliticized or not), often called "small words" by the unsophisticated consultant, have the pesky habit of appearing in abundance in texts. They may also show up in elicited sentences, but this cannot be relied upon. The fieldworker should begin with sentences found in texts, and then substitute one particle for the other (or zero) in order to get to the meaning. Asking about combinations and particle orders and varying their position in the sentence is also useful, but only if there is evidence from texts that particles can occur in combination with each other, and can have a variable position in the sentence.

One important subtype of ancillary elicitation to be used for syntax is Paraphrasing elicitation. The fieldworker asks the consultant to paraphrase text or other materials gathered during previous sessions (Samarin 1967b:119). Like target language translation (Section 12.4.2.2), this procedure is always easy on the fieldworker, but sometimes hard on the consultant. In addition, the consultant might not always understand why a paraphrase is useful to the fieldworker. S/he might think the fieldworker thought s/he made a mistake and now wants him or her to correct that mistake.

Texts cannot be used to get a large number of constructions quickly. There is just too much to look through in large corpora, and small text corpora will not show the full range of syntactic constructions. But as concordance programs and other computer software for language documentation become more and more sophisticated, doing computer searches can speed up the syntactic study of a corpus of texts. However, the fieldworker should always remember that a computer can only find the particular form or construction that the fieldworker asks for and has previously coded. Sometimes eyeballing hundreds of pages of text will reveal unusual or unexpected constructions that a computer was not instructed to find.

Some fieldworkers are not interested in the time-consuming work of text collection; in this case, we suggest that s/he work through ancillary elicitation on languages where text collections already exist. This will give the fieldworker some useful and valuable material to analyze.

12.5 Some Concluding Remarks

In this chapter, we might well come across as strongly encouraging ancillary elicitation, and strongly discouraging reverse translation and introspective judgment elicitation. We hope the beginning fieldworker can consider our advice to be therapeutic rather than strict and legalistic.

To be sure, not all data stemming from reverse translation and grammaticality judgments are unreliable. These approaches might be useful for filling in gaps in the analysis and for developing a deeper and clearer understanding of the grammar. Conversely, collecting texts and looking for data within texts are not aimless butterfly collecting activities (Everett 2004). Rather, they are approaches to building dependable corpora that avoid many of the problems associated with other elicitation tasks. All methods and sources are of value (Rice 2001:240).

The important point is that the fieldworker should not embark on the more delicate and perilous tasks before s/he is ready for them. The developing knowledge of the fieldworker and the developing personal relationship between fieldworker and consultant need to be taken into consideration when selecting the appropriate grammar gathering method.

Another important point is to encourage collaboration between the native speaker linguist and the non-native speaker linguist, particularly with the more difficult tasks, as suggested by Ameka (2006:71–77, 101–101). It is noteworthy that both structuralists such as Nida (1981), and formalists such as Bjerre et al. (2008:160) have come to the same conclusion regarding collaboration between native and non-native speakers when investigating subtle points of syntax.

References

Abbi, Anvita. 2001. *A Manual of Linguistic Field Work and Indian Language Structures.* (Lincom Handbooks in Linguistics 17.) Munich: Lincom Europa

Abraham, Werner. 1991a. Modal particle research: the state of the art. *Multilingua* 10(1/2):9–15.

Abraham, Werner, ed. 1991b. *Discourse particles: descriptive and theoretical investigations in the logical, syntactic and pragmatic properties of discourse particles in German.* Amsterdam and Philadelphia: John Benjamins.

Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y. 2007. Typological Distinctions in Word-Formation. In *Language Typology and Syntactic Description. Vol. 3: Grammatical categories and the lexicon*, ed. by Timothy Shopen, 1–65. Cambridge: Cambridge University Press. [2nd edition].

Aitken, Barbara. 1955. A note on eliciting. *International Journal of American Linguistics* 21(1):81.

Ameka, Felix K. 2006. Real descriptions: Reflections on native speaker and non-native speaker descriptions of a language. In *Catching language: The Standing Challenge of Grammar Writing*, ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 69–112. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Ammon, U., N. Dittmar, and K. Mattheier, eds. 1988. *Sociolinguistics* Vol. 2. Chapter 8, Elicitation methods. Berlin: Mouton de Gruyter.

Anttila, Raimo. 1972. *An Introduction to Historical and Comparative Linguistics*. New York: Macmillan.

Badten, Linda Womkon, Vera Oovi Kaneshiro, Marie Oovi, Christopher Koonooka, and Steven A. Jacobson. 2008. *St. Lawrence Island/Siberian Yupik Eskimo Dictionary*. Fairbanks: Alaska Native Language Center, University of Alaska.

Barðdal, Jóhanna, and Shobhana Chelliah, eds. 2009. *The Role of Semantic, Pragmatic and Discourse Factors in the Development of Case*. Amsterdam and Philadelphia: John Benjamins.

Barnum, Francis, S. J. 1893. *Life on the Alaska Mission*. Woodstock, MD: Woodstock College Press.

Basset, André. 1951. L'enquête linguistique. *Conférence de l'Institut de Linguistique de l'Université de Paris* X. Années 1950–1951, 7–22. Paris: Librairie C. Klincksieck.

Berlin, Brent, and Terrence Kaufman. 1985–1987. *South American Indian Languages Documentation Project* (*SAILP). Projeto de Documentação das Linguas da America del Sul.* Introdução, Questionario Gramatical, Questionario Lexical. Department of Anthropology, University of California, Berkeley, and Department of Linguistics, University of Pittsburgh.

Bjerre, Tavs, Eva Engels, Henrik Jørgensen, and Sten Vikner. 2008. Points of convergence between functional and formal approaches to syntactic analysis. *Working Papers in Scandinavian Syntax* 82:131–166. Lund: Center for Language and Literature.

Bliss, Heather, and Elizabeth Ritter. 2001. Developing a Database of Personal and Demonstrative Pronoun Paradigms: Conceptual and Technical Challenges. In *Proceedings of the ICRS Workshop on Linguistic Databases*, ed. by Steven Bird, Peter Buneman, and Mark Lieberman.

Bohnemeyer, Jürgen, Melissa Bowerman, and Penelope Brown. 2001. Cut and break clips. In *Manual for the Field season 2001*, ed. by Stephen C. Levinson and Nicholas J. Enfield. Nijmegen: Max Planck Institute Language and Cognition Group.

Bouquiaux, Luc, Jacqueline M. C. Thomas, and James Robert. 1992. *Studying and Describing Unwritten Languages.* Dallas, TX: SIL International.

Bowen, Elenore Smith. 1964. *Return to Laughter*. 2nd ed. New York: Doubleday.

Bowern, Claire. 2008. *Linguistic Fieldwork. A Practical Guide*. New York: Palgrave MacMillan.

Bridgeman.Lorraine I, and Sarah Gudschinsky. 1965. Plano para pesquisas nas línguas Tupi. In O Setor Linguistico do Museu Nacional (Organização e objetivos), ed. by L. de Castro Faria, 17–25. *Publicações avulsas do Museu Nacional* No. 49. Rio de Janeiro: Universidade Federal do Rio de Janeiro.

Bryant Margaret. 1945. *Proverbs and How to collect Them*. Greensboro: NC: American Dialect Society.

Capell, Arthur. 1965. *Linguistic materials for fieldworkers in Australia*. Canberra: Australian Institute of Aboriginal Studies.

Carroll, John, Thomas Bever, and Chava Pollack. 1981. The Non-Uniqueness of Linguistic Intuitions. *Language* 57(2):368–383.

Chafe, Wallace L., ed. 1980. *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. Norwood, NJ: Ablex.

Chelliah Shobhana L. 1997. *A Grammar of Meithei*. (Mouton Grammar Library 17.) Berlin and New York: Mouton de Gruyter.

Chelliah, Shobhana L. 2001. The role of text collection and elicitation in linguistic fieldwork. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 152–165. Cambridge: Cambridge University Press.

Chomsky, Noam A. 1957. *Syntactic Structures*. The Hague: Mouton.

Cohen Marcel S. 1950-1951. *Questionnaire Linguistique*. Nimègue, The Netherlands: Comité International Permanent de Linguistes, Commission d'Enquête Linguistique.

Comrie, Bernard, and Norval Smith. 1977. Lingua Descriptive Studies: Questionnaire. *Lingua* 42.1–72.

Coulmas, Florian, ed. 1981a. *A Festschrift for Native Speaker*. The Hague, Paris and New York: Mouton.

Coulmas, Florian. 1981b. Introduction: The Concept of Native Speaker. In *A Festschrift for Native Speaker*, ed. by Florian A. Coulmas, 1–25. The Hague, Paris and New York: Mouton.

Cowart, Wayne. 1997. *Experimental Syntax: Applying Objective Methods to Sentence Judgments*. Thousand Oaks, CA: Sage Publications.

Crain, Stephen, and Rosalind Thornton. 1998. *Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax*. Cambridge, MA: MIT Press.

Crowley, Terry. 2007. *Field Linguistics. A Beginner's Guide*. Oxford: Oxford University Press.

Dale, Ian R. H. 1978. Beyond intuition: the use of questionnaires in linguistic investigation. *Anthropological Linguistics* 20(4):158–66.

Diller, Anthony V. N., and Wilaiwan Khanittanan. 2002. Syntactic Enquiry as a Cultural Activity. In *Ethnosyntax. Explorations in Grammar and Culture*, ed. by Nicholas J. Enfield, 30–51. Oxford and New York: Oxford University Press.,

Dimmendaal, Gerrit J. 2001. Places and People: fieldsites and informants. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 55–75. Cambridge: Cambridge University Press.

Dixon, R. M. W. 2007. Field Linguistics: a minor manual. *Sprachtypologie und Universalienforschung* (*Focus on Linguistic Fieldwork*, ed. by Alexandra Aikhenvald) 60(1).12–31.

Dixon, R. M. W. 2010a. *Basic Linguistic Theory. Volume 1 Methodology*. Oxford: Oxford University Press.

Dixon, R. M. W. 2010b. *Basic Linguistic Theory. Volume 2 Grammatical Topics*. Oxford: Oxford University Press.

Dooley Robert, A. 1989. Suggestions for the field linguist regarding quotations. *Notes on Linguistics* 44:34–50.

Dorian, Nancy C. 1986. Gathering Language data in terminal speech communities. In *The Fergusonian Impact*, *In honor of Charles A. Ferguson on the occasion of his 65th birthday*. Vol. 2: Sociolinguistics and the sociology of language, ed. by Joshua A. Fishman, Andree Tabouret-Keller, Michael Clyne, Bh. Krishnamurti, and M. Abdulaziz, 555–575. Berlin: Mouton de Gruyter.

Dorian, Nancy C. 2001. Surprises in Sutherland: linguistic variability amidst social uniformity. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 133–151. Cambridge: Cambridge University Press.

Enfield, Nicholas J. 2002. Cultural Logic and Syntactic Productivity: Associated Posture Constructions in Lao. In *Ethnosyntax. Explorations in Grammar and Culture*, ed. by Nicholas J. Enfield, 231–258. Oxford and New York: Oxford University Press.

Enfield, Nicholas J., and Michael Dunn. 2001. *Supplements to the Wilkins 1999 Demonstrative Questionnaire*, 82-84. Nijmegen: Max Planck Institute Language and Cognition Group.

van Engelenhoven, Aone. 2000. *Deixis and Location Questionnaire*. 2000 East Nusantara Linguistics Workshop, Deixis Workshop, Australian National University. Online: http://rspas.anu.edu/au/linguistics/projects/Conferences/EastNusantara/DeixisQuest.html

Everett, Daniel L. 2001. Monolingual field research. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 166–188. Cambridge: Cambridge University Press.

Everett, Daniel L. 2004. Coherent fieldwork. In *Linguistics today—Facing a greater challenge*, ed. by Piet van Sterkenburg, 141–162. Amsterdam and Philadelphia: John Benjamins.

Foley, William A. 2002. Field methods. In *The Linguistics Encyclopedia*. Second ed., ed. by Kirsten Malmkjær, 131–137. London and New York: Routledge.

Freeze, Ray A. 1989. *Mayo de los Capomos, Sinaloa*. (Archivo de Lenguas Indígenas de México.) Mexico City: El Colegio de México.

García, Erica C. 1967. Auxiliaries and the criterion of simplicity. *Language* 43(4):853–870.

Geniušienė Emma. 1987. The Typology of Reflexives. (Empirical Approaches to language Typology 2.) Berlin: Mouton de Gruyter.

Gil, David. 1982. *Distributive Numerals*. Ph.D. diss. University of California, Los Angeles.

Gil, David. 2001. Escaping Eurocentrism: fieldwork as a process of unlearning. In *Linguistic fieldwork*, ed. by Paul Newman and Martha Ratliff, 102–132. Cambridge: Cambridge University Press.

Grimes, Joseph E. 1967. Positional analysis. *Language* 43(2):437–444.

Greenbaum, Sidney. 1973. Informant Elicitation of Data on Syntactic Variation. *Lingua* 31:201–212.

Greenbaum, Sidney. 1976. Syntactic Frequency and Acceptability. *Lingua* 40(2/3):99–113.

Greenbaum, Sidney. 1977. Judgements of Syntactic Acceptability. *Studia Linguistica* 31(2):83–105.

Hale, Kenneth L. 1965. On the use of informants in fieldwork. *Canadian Journal of Linguistics* 10:108–119.

Hale, Kenneth L. 2001. Ulwa (Southern Sumu) Places and People: the beginnings of a language research project. In *Linguistic fieldwork*, ed. by Paul Newman and Martha Ratliff, 76–101. Cambridge: Cambridge University Press.

Harris, Zellig S. and Charles F. Voegelin. 1953. Eliciting in Linguistics. *Southwestern Journal of Anthropology* 9:59–75.

Haviland, John B. 2006. Documenting lexical knowledge. In *Essentials of Language Documentation*, ed. by Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel, 129–162. (Trends in Linguistics. Studies and Monographs 178.) Berlin and New York: Mouton de Gruyter.

Hayes, Alfred S. 1954. Field procedures while working with Diegueño. *International Journal of American Linguistics* 20:185–94.

Healey, Alan. 1964. *Handling Unsophisticated Linguistic Informants*. (Linguistic Circle of Canberra Publications, Series A, Occasional Papers No. 2.) Canberra: Linguistics, Department of Anthropology and Sociology, The Australian National University.

Healey, Alan, ed. 1975. *Language Learner's Field Guide*. Ukarumpa, Papua New Guinea: Summer Institute of Linguistics.

Hellwig, Birgit. 2006. Field semantics and grammar-writing: Stimuli-based techniques and the study of locative verbs. In *Catching language: The Standing Challenge of Grammar Writing*.), ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 321–358. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Hellwig, Birgit. 2007. Fieldwork among the Goemai in Nigeria: discovering the grammar of property expressions. *Sprachtypologie und Universalienforschung* (Focus on Linguistic Fieldwork, ed. by Alexandra Aikhenvald) 60(1):67–80.

Hensel, Chase, Marie Blanchett, Ida Alexie, and Phyllis Morrow. 1983. *Qaneryaurci Yup'igtun*. Bethel, AK: Yup'ik Language Center.

Himmelmann, Nikolaus P. 2006. The challenges of segmenting spoken language. In *Essentials of Language Documentation*, ed. by Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel, 253–274. (Trends in Linguistics. Studies and Monographs 178.) Berlin and New York: Mouton de Gruyter.

Hinds, John. 1981. The interpretation of Ungrammatical Utterances. In *A Festschrift for Native Speaker*, ed. by Florian A. Coulmas, 221–235. The Hague, Paris and New York: Mouton.

Hinton, Leanne, Johanna Nichols, and John Ohala, eds. 1994. *Sound Symbolism*. Cambridge: Cambridge University Press.

Hinton, Leanne, Matt Vera, Nancy Steele, and the Advocates for Indigenous Language Survival. 2002. *How to Keep Your Language Alive: A Common Sense Approach to One-on-One Language Learning*. Berkeley, CA: Heyday Books.

Hoijer, Harry. 1938. *Chiricahua and Mescalero Apache Texts*. Chicago: The University of Chicago Press.

Hoijer, Harry. 1946. Chiricahua Apache. In *Linguistic Structures of Native America*, ed. by Cornelius Osgood, 55–84. (Viking Fund Publications in Anthropology 6.) New York: Wenner Gren.

Hopkins, Jill D., and Louanna Furbee. 1991. Indirectness in the interview. *Journal of Linguistic Anthropology* 1(1):63–77.

Hyman, Larry M. 2001. Fieldwork as a state of mind. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 15-33. Cambridge: Cambridge University Press.

Itkonen, Esa. 1981. The Concept of Linguistic Intuition. In *A Festschrift for Native Speaker*, ed. by Florian A. Coulmas, 127–140. The Hague, Paris and New York: Mouton.

Johnston, Raymond L. 1980. *Grammar and Basic Vocabulary in Oceanic Austronesian Languages: A Standard Elicitation Schedule. Ukarumpa*, Papua New Guinea: Summer Institute of Linguistics.

Johnstone, Barbara. 2000. *Qualitative methods in Sociolinguistics*. New York: Oxford University Press.

Kallionen, Vilho. 1974. *Finnish Conversational Exercises. Elementary Level*. Helsinki: Suomalaisen Kirjalisuuden Seura.

Kaufman, Terrence. 1990. Language History in South America: What we Know and How to Know More. In *Amazonian Linguistics. Studies in Lowland South American Languages*, ed. by Doris L. Payne, 13–73. Austin: University of Texas Press.

Kaufman, Terrence, with the assistance of John Justeson 2003. *A Preliminary Mayan Etymological Dictionary*. University of Pittsburgh, Department of Linguistics. Online: http://www.famsi.org/reports/01051/pmed.pdf

Kibrik, Aleksandr E. 1977. *The Methodology of Field Investigations in Linguistics*. (Setting up the Problem.) (Janua Linguarum, Series Minor, 142.) The Hague: Mouton.

Kirschenblatt-Gimblett, Barbara, ed. 1976. *Speech Play. Research and Resources for Studying Linguistic Creativity*. Philadelphia: University of Pennsylvania Press.

Klamer, Marian. 2000. *Valence Questionnaire*. 2000 East Nusantara Linguistics Workshop, Valency Workshop, Australian National University. Online: http://rspas.anu.edu/au/linguistics/projects/Conferences/EastNusantara/ValenceQuest.html

König, Ekkehard, (with Dik Bakker, Östen Dahl, Martin Haspelmath, Maria Koptjevskaja-Tamm, Christian Lehmann, and Anna Siewierska). 1993. *EUROTYP Guidelines*. Berlin and Strasbourg: European Science Foundation Programme in Language Typology. Online: http://www-uilots.let.uu.nl/ltrc/eurotyp/h5.htm

Krauss, Michael E. 2006. A History of Eyak Language Documentation and Study: Fredericae de Laguna in Memoriam. *Arctic Anthropology* 43(2):172–218.

Krishnamurti, Bhadriraju. 2007. Fieldwork on Konda, a Dravidian language. *Sprachtypologie und Universalienforschung* (Focus on Linguistic Fieldwork, ed. by Alexandra Aikhenvald) 60(1):56–66.

Labov, William. 1972. Some principles of linguistic methodology. *Language in Society* 1:97–120.

Labov, William. 1973. The Place of Linguistic Research in American Society. In *Themes in Linguistics: The 1970s*, ed. by Eric P. Hamp, 97–129. The Hague: Mouton.

Labov, William. 1975. Empirical Foundations of Linguistic Theory. *The Scope of American Linguistics. The First Golden Anniversary Symposium of the Linguistic Society of America*, ed. by Robert Austerlitz, 77–133. Lisse: The Peter de Ridder Press.

Labov, William. 1996. When intuitions fail. In *Papers from the Parasession on Theory and Data in Linguistics*, *Chicago Linguistic Society* 32:77–105, ed. by Lisa McNair. Chicago: University of Chicago Press.

Lastra de Suárez, Yolanda. 1986. *Las áreas dialectales del náhuatl moderno*. México, D.F.: Universidad Nacional Autónoma de México.

Laughlin, Robert M. 1975. *The Great Tzotzil Dictionary of San Lorenzo Zinacantán*. (Smithsonian Contributions to Anthropology No. 19.) Washington D.C.: Smithsonian Institution Press.

Leaders, Marlin. 1991. Eliciting figures of speech. *Notes on Translation* 5(4):31–45.

Levinson, Stephen C. 1992. Primer for the field investigation of spatial description and conception. *Pragmatics* 2(1):5–47.

Longacre, Robert E. 1964. *Grammar Discovery Procedures: A Field Manual*. The Hague: Mouton and Co.

Maclay H., and M. D. Sleator. 1960. Responses to language: judgments of grammaticalness. *International Journal of American Linguistics* 26:275–82.

McDavid, Raven. 1985. Eliciting: Direct, Indirect, and Oblique. *American Speech* 57:228–231.

Maddieson, Ian. 2001. Phonetic fieldwork. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 211-229. Cambridge: Cambridge University Press.

Matthewson, Lisa. 2004. On the methodology of semantic fieldwork. *International Journal of American Linguistics* 70(4):369–415.

Max Planck Institute for Evolutionary Anthropology. 2007 *A Typological Approach to Field Linguistics. Tools for Language Description*. Max Planck Institute for Evolutionary Anthropology, Dept. of Linguistics, Leipzig, Germany. Online: http://lingweb.eva.mpg.de/fieldtools/tools.htm

Mayer, Mercer. 1969. *Frog, Where are You?* New York: Dial Books for Young Readers.

Maynor, Natalie. 1982. Grammatical Judgments in LANE. *American Speech* 57:228–231.

Milroy, Lesley. 1987. *Observing and Analysing Natural Language. A Critical Account of Sociolinguistic method*. (Language in Society 12.) Oxford: Basil Blackwell.

Mintz, Malcolm Warren. 1986. *Kursus Perbualan Bahasa Malaysia. A course in conversational Malay*. Perth, Western Australia: School of Human Communication: Murdoch University.

Mithun, Marianne. 1990. Language obsolescence and grammatical description. *International Journal of American Linguistics* 56(1):1–26.

Mithun, Marianne. 2001. Who shapes the record: the speaker and the linguist. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 34–54. Cambridge: Cambridge University Press.

Mithun, Marianne. 2007a. Grammars and the Community. In *Perspectives on Grammar Writing*, ed. by Thomas E. Payne and David J. Weber, 45–69. (Benjamins Current Topics 11.) Amsterdam and Philadelphia: John Benjamins.

Mithun, Marianne. 2007b. Linguistics in the face of language endangerment. In *Language Endangerment and Endangered languages. Linguistic and Anthropological Studies with Special Emphasis on the Languages and Cultures of the Andean-Amazonian Border Area*. (Indigenous Languages of Latin America 5.), ed. by W. Leo Wetzels, Leiden: CNWS. 15–34.

Mosel, Ulrike. 2006a. Grammaticography: The art and craft of writing grammars. In *Catching language: The Standing Challenge of Grammar Writing*, ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 11–68. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton de Gruyter.

Mosel, Ulrike. 2006b. Fieldwork and community language work. In *Essentials of Language Documentation*, ed. by Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel, 67–85. (Trends in Linguistics. Studies and Monographs 178.) Berlin and New York: Mouton de Gruyter.

Munro, Pamela. 2003. Field Linguistics. In *The Handbook of Linguistics*, ed. by Mark Aronoff and Janie Rees-Miller, 130–149. Malden, MA: Blackwell Publishing.

Newman, Paul, and Martha Ratliff. 2001. Introduction. In *Linguistic fieldwork*, ed. by Paul Newman and Martha Ratliff, 1–14. Cambridge: Cambridge University Press.

Newmeyer, Frederick J. 1983. *Grammatical Theory. Its Limits and Its Possibilities*. Chicago and London: University of Chicago Press.

Nida, Eugene A. 1947. Field Methods in Descriptive Linguistics. *International Journal of American Linguistics* 13:138–146.

Nida, Eugene A. 1950. Training the translation helper. *The Bible Translator* 1:56–62.

Nida, Eugene A. 1952-53. Selective listening. *Language Learning* 4:92–101.

Nida, Eugene A. 1981. Informants or Colleagues? In *A Festschrift for Native Speaker*, ed. by Florian A. Coulmas, 169–174. The Hague, Paris and New York: Mouton.

Noonan, Michael. 1985. Complementation. In *Language Typology and Syntactic Description. Vol. 2: Complex constructions*, ed. by Timothy Shopen, 42–140. Cambridge: Cambridge University Press.

Nuckolls, Janis B. 1996. *Sounds like Life. Sound-symbolic Grammar, Performance, and Cognition in Pastaza Quechua*. (Oxford Studies in Anthropological Linguistics 2.) New York and Oxford: Oxford University Press.

O'Connor, Loretta. 2007. *Motion, Transfer and Transformation. The grammar of change in Lowland Chontal*. Amsterdam and Philadelphia: John Benjamins.

Paikeday, Thomas M. 1985. *The Native Speaker is Dead! An informal discussion of a linguistic myth with Noam Chomsky and other linguists, philosophers, psychologists, and lexicographers*. Toronto and New York: Lexicography, Inc. (formerly Paikeday Publishing Inc.). [Online edition 2003, http://www.paikeday.net/speaker.pdf]

Palmer, Adrian S. 1974. *AUA Language Center Thai Course, Dialog Book A: Small Talk*. Bangkok: The American University Alumni Association.

Paul, Benjamin D. 1953. Interview techniques and field relationships. In *Anthropology Today: Selections*, ed. by Sol Tax. Chicago: University of Chicago Press.

Payne, Thomas E. 1997. *Describing morphosyntax: a guide for field linguists*. Cambridge: Cambridge University Press.

Plank, Frans, ed. 1995. *Double case: Agreement by Suffixaufname*. New York: Oxford University Press.

Pop, Sever. 1955. *Bibliographie des Questionnaries Linguistiques*. Louvain: Commission d'Enquête Linguistique, Publication 6.

Prost, André. 1974. Plan Pour L'Etude et La Description d'une Langue. In *Les langues sans tradition écrite; méthodes d'enquête et de description*. (Actes du Colloque International du CNRS, Nice 28 Juin-2 Juillet 1971.), 75–89. Paris: Société d'Études Linguistiques et Anthropologiques de France. No. 3, Numéro spécial.

de Reuse, Willem J. 1994. *Siberian Yupik Eskimo. The language and its contacts with Chukchi*. Salt Lake City: University of Utah Press.

Reynolds, Jon. 1993. Difficulty of the Native Language. *Georgetown*, Winter 1993: 62.

Rice, Keren. 2001. Learning as one goes. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 230–249. Cambridge: Cambridge University Press.

Richards, I. A., David Weinstein, and Christine Gibson. 1961. *Hebrew Through Pictures*, Book 1. New York: Washington Square Press.

Ringen, Jon. 1977. On evaluating data concerning linguistic intuition. In *Current Themes in Linguistics*, ed. by Fred Eckman, 145–162. Washington DC: Hemisphere Publishing.

Ringen, Jon. 1981. Quine on Introspection in Linguistics. In *A Festschrift for Native Speaker*, ed. by Florian A. Coulmas, 141–151. The Hague, Paris and New York: Mouton.

Rood, David S. 2003. Two Lakhota Locatives and the Role of Introspection in Linguistic Analysis. In *Motion, Direction and Location in Languages*, ed. by Erin Shay and Uwe Seibert, 255–258. Amsterdam and Philadelphia: John Benjamins.

Ross, John R. 1979. Where's English? In *Individual differences in language ability and language behavior*, ed. by Charles J. Fillmore, W. Kempler and W. S.-Y. Wang, 127–165. New York: Academic Press.

Rupp, James. 1974. On eliciting metaphors. *Notes on Translation* 53:17–22.

Samarin, William J. 1967a. Determining the Meanings of Ideophones. *Journal of West African Languages* 6(2):35–41.

Samarin, William J. 1967b. *Field Linguistics. A Guide to Linguistic Field Work*. New York: Holt, Rinehart and Winston.

Samarin, William J. 1970. Field Procedures in Ideophone Research. *Journal of African languages* 19:27–30.

Samarin, William J. 1974. Semantics Without Native Intuition. In *Les langues sans tradition écrite; méthodes d'enquête et de description*. (Actes du Colloque International du CNRS, Nice 28 Juin-2 Juillet 1971.), 159–174. Paris: Société d'Études Linguistiques et Anthropologiques de France. No. 3, Numéro spécial.

Samarin, William J. 1991. Intersubjective and Interdialectal Variation in Gbeya Ideophones. *Journal of Linguistic Anthropology* 1(1):52–62.

Sampson, Geoffrey. 2002. *Empirical Linguistics*. London and New York: Continuum.

Sandalo, Filomena. 1997. *A Grammar of Kadiweu, with special reference to the Polysynthesis Parameter*. (MIT Occasional papers in Linguistics No. 11.) Cambridge, MA: MIT.

Sapon, Stanley M. 1957. *A pictorial linguistic interview manual*. (American Library of Recorded Dialect Studies.) Columbus: Ohio State University.

Schneider, Hans Julius. 1995. Intuition and Introspection. In *Handbook of Pragmatics: Manual*, ed. by Jef Verschueren, Jan-Ola Ostman, and Jan Blommaert, 606–608. Amsterdam and Philadelphia: John Benjamins.

Schnelle, H. 1981. Introspection and the Description of Language Use. In *A Festschrift for Native Speaker*, ed. by Florian A. Coulmas, 105–126. The Hague, Paris and New York: Mouton.

Schwarte, Barbara. 1974. Intuitions of grammaticality and the 'law of contrast': A pilot study. *Studies in the Linguistic Sciences* 4(1):198–216.

Seiler, Hansjakob. 1970. *Cahuilla Texts with an Introduction*. Bloomington: Indiana University.

Senft, Gunter. 1995. Elicitation. In *Handbook of Pragmatics: Manual*, ed. by Jef Verschueren, Jan-Ola Ostman, and Jan Blommaert, 577–581. Amsterdam and Philadelphia: John Benjamins.

Sherzer, Joel. 1982. Play Languages: With a Note on Ritual Languages. In *Exceptional language and Linguistics*, ed. by Lorraine Obler and Lise Menn, 175–99. New York: Academic Press.

Sherzer, Joel. 2002. *Speech Play and Verbal Art*. Austin: University of Texas Press.

Shopen, Timothy, ed. 1985. *Language Typology and Syntactic Description. Vol. 1: Clause structure; Vol. 2: Complex constructions; Vol. 3: Grammatical categories and the lexicon*. Cambridge: Cambridge University Press.

Siewierska, Anna. 1993. EUROTYP Word Order Questionnaire. Ms, 25 p. Online: http://www.eva.mpg.de/lingua/tools-at-lingboard/pdf/siewierska.pdf.

Silverstein, Michael. 1979. Language Structure and Linguistic Ideology. In *The Elements: A Parasession in Linguistic Units and Levels*, *Chicago Linguistic Society*, ed. by R. Clyne, William Hanks and Carol Hofbauer, 193–247. Chicago: Chicago Linguistic Society.

Spencer, N. 1973. Differences Between Linguists and Nonlinguists in Intuitions of Grammaticality-Acceptability. *Journal of Psycholinguistic Research* 2(2):83–98.

Summer Institute of Linguistics. 1965. Formulário para estudos comparativos nas línguas Brasileiras. In O Setor Linguistico do Museu Nacional (Organização e objetivos), ed. by

L. de Castro Faria, 27–36. *Publicações avulsas do Museu Nacional* No. 49. Rio de Janeiro: Universidade Federal do Rio de Janeiro.

Sutton, Peter, and Michael Walsh. 1979. *Revised linguistic fieldwork manual for Australia*. Canberra: Australian Institute of Aboriginal Studies.

Swerts M. and R. Collier. 1992. On the controlled elicitation of spontaneous speech. *Speech Communication* 11(4/5):463–68.

Talmy, Leonard. n. d. *Introspection as a Methodology in Linguistics*. Handout, Department of Linguistics, Center for Cognitive Science, State University of New York at Buffalo. Online: http://linguistics.buffalo.edu/people/faculty/talmy/talmyweb/Handouts/Introspection2.pdf

Toulouse, Isadore. 1995. *Kidwenan. An Ojibwe Language Book*. Revised Edition. Munsee-Delaware Nation, Ontario: Anishinaabwe Kendaaswin Publishing.

Tucker, A. N. 1974. Questionnaire pour les langues Bantu et Non-Bantu. In *Les langues sans tradition écrite; méthodes d'enquête et de description*. (Actes du Colloque International du CNRS, Nice 28 Juin-2 Juillet 1971.), 49–74. Paris: Société d'Études Linguistiques et Anthropologiques de France. No. 3, Numéro spécial.

Ulvestad, Bjarne. 1981. On the Precariousness of Linguistic Introspection. In *A Festschrift for Native Speaker*, ed. by Florian A. Coulmas, 245–261. The Hague, Paris and New York: Mouton.

Vandeloise, Claude. 1991. *Spatial Prepositions. A Case Study from French*. Chicago and London: University of Chicago Press.

Vaux, Bert, Justin Cooper, and Emily Tucker. 2007. *Linguistic Field Methods*. Eugene, OR: Wipf and Stock.

Voegelin, Carl F. 1959a. Review of Eastern Ojibwa, by Leonard Bloomfield. *Language* 35(1):109–125.

Voegelin, Carl F. 1959b. The notion of arbitrariness in structural statement and restatement I: Eliciting. *International Journal of American Linguistics* 25(4):207–20.

Voegelin, Carl F. 1960. Subsystem typology in linguistics. In *Man and Cultures: Selected papers*, ed. by Anthony F. C. Wallace, 202–206. (Fifth International Congress of Anthropological and Ethnological Sciences). Philadelphia: University of Pennsylvania Press.

Wetzels, W. Leo, ed. 2007. *Language Endangerment and Endangered languages. Linguistic and Anthropological Studies with Special Emphasis on the Languages and Cultures of the Andean-Amazonian Border Area*. (Indigenous Languages of Latin America 5.) Leiden: CNWS.

Whiteley, William H. 1964. Suggestions for recording a Bantu language in the field. *Tanganyika Notes and Records*, No. 62:1–19.

Wilkins, David. 1999. Demonstratives questionnaire: "THIS and "THAT" in comparative perspective. In *"Manual" for the 1999 field season*, ed. by David Wilkins, 1–24. Nijmegen: Max Planck Institute Language and Cognition Group.

Wilson, Alan. 1969. *Breakthrough Navajo. An Introductory Course*. Gallup, NM: University of New Mexico, Gallup Branch.

Wilson, Alan, and Gene Dennison. 1978. *Speak Navajo. An Intermediate Text in communication*. Gallup, NM: University of New Mexico, Gallup Branch.

Wilson, Darryl. 1990. Idiom Discovery procedure. *Notes On Linguistics* 49:4–6.

Wolfram, Walt. 1986. Good data in a bad situation: Eliciting Vernacular Structures. In *The Fergusonian Impact*, *In honor of Charles A. Ferguson on the occasion of his 65th birthday*. Vol. 2: Sociolinguistics and the sociology of language, ed. by Joshua A. Fishman, Andree Tabouret-Keller, Michael Clyne, Bh. Krishnamurti, and M. Abdulaziz, 3–22. Berlin: Mouton de Gruyter.

Wolfram, Walt, and Ralph Fasold. 1974. *The Study of Social Dialects in American English*. Englewood Cliffs, NJ: Prentice-Hall.

Wong, I. F. H. 1975. Field Procedures in Generative Grammar. *Anthropological Linguistics* 17(2):43–52.

Woodbury, Anthony C. 1993. A defense of the proposition: "When a language dies, culture dies." (Proceedings of the First Annual Symposium About Language and Society-Austin (SALSA)) *Texas Linguistic Forum* 33:101–129.

Wunderlich, Dieter. 1981. Linguistic Strategies. In *A Festschrift for Native Speaker*, ed. by Florian A. Coulmas, 279–296. The Hague, Paris and New York: Mouton.

Yegerlehner, J. 1955. A note on eliciting techniques. *International Journal of American Linguistics* 21:286–88.

Chapter 13
Semantics, Pragmatics, and Text Collection

"…the search for meaning in any language is best seen as a never-ending stringing together of hypertextual commentary which gradually leads to a better understanding of the utterances under study" (Evans and Sasse 2007:260)

13.1 Introduction

In this final chapter, we discuss the importance of semantics and pragmatics to descriptive linguistic fieldwork. We also provide a guide to text collection, which we feel should be a central part of the corpus used for language analysis, and which is intimately linked to issues of semantics and pragmatics.

13.2 Semantics and Fieldwork

In language description, the meaning of utterances must be adequately established to conduct lexical, morphological, and syntactic analysis. Therefore, semantics, the study of meaning, is an integral part of descriptive linguistic field projects. In this section, we discuss concepts and terminology from subfields of semantics which can be useful in the investigation of other aspects of linguistic inquiry and analysis.

13.2.1 Lexical Semantics

As part of a language description or documentation project, the fieldworker may want to investigate how the words of a language are related. One way to do this is to find lexical fields which group words together. This grouping can be based on semantic similarity; for example, a possible lexical field might be a set of color terms or a set of kinship terms. Another way to organize word groups is hierarchically, identifying a superordinate term (a hyponym) and related subordinate terms

S.L. Chelliah and W.J. de Reuse, *Handbook of Descriptive Linguistic Fieldwork*,
DOI 10.1007/978-90-481-9026-3_13, © Springer Science+Business Media B.V. 2011

(hypernyms); for example, superordinate *bird*, and subordinate *robin, pigeon, eagle,* and *peacock*. Such word nets are useful for semantic decomposition and morphological analysis. When semantically similar words are grouped together in word nets, the analyst may find the same morphology or word structure repeated in the grouped words. The meaning of the morpheme can be more easily arrived at because there will be multiple examples from which to extrapolate.

Other word relationships of interest are synonomy, homophony, antonymy, and polysemy. Homophony and polysemy, for instance, become important when glossing morphemes and compiling word lists or dictionaries. Consider this set of examples of the English word *run* from Saeed (2009:60).

1. I go for a *run* every morning.
2. The tail-end batsmen added a single *run* before lunch.
3. The ball-player hit a home *run*.
4. We took the new car for a *run*.
5. He built a new *run* for his chickens.
6. There's been a *run* on the dollar.
7. The bears are here for the salmon *run*.

Are all the instances of *run* in 1–7 different forms that are accidentally homophonous? If so, the forms should be listed as separate entries. Alternatively, if all the instances of *run* in 1–7 have the same core semantic representation but different related senses which speakers interpret based on particular contexts of use, the forms are polysemous.[1] Polysemous examples of *run* warrant a single lexical entry and the same gloss with explanations of meaning differences. In the case of polysemy, decisions must be made about appropriate glossing of morphemes. In Manipuri (Tibeto-Burman, Northeast India), for example, the same consonant and vowel sequence is used in different but semantically related ways: *yá* can mean '(be) possible, (be) able, (be) near, (to) agree, (to) accept, (to) yield'. Which meaning is basic? Should the gloss be uniform or reflect the different meanings?

Similar questions are posed by Enfield (2006:297) about a particular type of polysemy called heterosemy. In this type of semantic relationship, lexical items are related in meaning but belong to different lexical categories. Examples are *father*, *stone*, *talk,* and *walk* in English, where both verb and noun forms are attested. For grammatical description – specifically for text and example annotation – the question is whether to list such related forms as separate entities, or else as the same entity, with each item derivable by rule. A related lexicographical issue discussed in Evans and Sasse (2007:267) is that it is impossible to elicit these multiple meanings from single prompts such as pictures or videos. A picture of a stone will usually only elicit the noun.

Synonyms are words that have very similar meanings but almost always have different connotations. Glossing should take this into consideration. Best practice requires that different forms not be glossed the same way (Comrie et al. 2008).

[1] We avoid using the terms 'lexical ambiguity' (often used to describe cases of homophony) and 'lexical vagueness' as these are not particularly useful for lexicographic or annotation tasks.

If the variants of a synonym are used in different registers, a convention will need to be adopted to make this clear; for instance, in providing English glosses, a Latinate form could be given for the word from a higher or more formal register, and a Germanic equivalent given for the lower or more informal register (Franchetto 2006).

Antonyms can be used to elicit new lexical items or to determine the meaning of elicited items. It should be kept in mind that not all antonymic relationships are the same. Saeed (2009:66–69) identifies the following possibilities:

- Simple antonyms: "the negative of one implies the positive of another" (67)
- Gradable antonyms: words describe points on a continuum such as *hot*, *warm*, and *cold* (67)
- Reverse antonyms: word pairs describe opposite actions or movements, such as *push* and *pull*, or *deflate* and *inflate*
- Converse antonyms: word pairs reflect opposite points of view, such as *employer* and *employee,* or *buy* and *sell*
- Taxonomic sisters: terms in the same lexical field refer to different things, e.g., color terms, or words for ice cream flavors

Concepts from lexical semantics also inform the glossing of compounds. In many cases, speakers cannot access the meanings of individual stems in compounds. This can be frustrating for the linguist to whom the meaning of individual stems may seem fairly obvious. An example is Manipuri *yumgom* 'outhouse, urinal' which I [Chelliah] have analyzed as *yum-khom,* house-collect. Not all speakers agree that the second stem is the same as *khom* 'collect', presumably because a phonological rule obscures the underlying form or the stem. Even more serious problems can arise when speakers are prone to literal rather than figurative interpretation of speech (see also Section 11.2.2.4). For these speakers, it will be difficult to explain compounds whose meanings are based on metaphor or hyperbole, so elicitation will have to be supplemented with historical and morphological analysis to gloss their stems.

Words which are related as parts of a whole belong to meronymic hierarchies (Saeed 2009:71). Such hierarchies can be used to elicit new words by finding an entity or construction with many components, and then eliciting names for those components or parts. The task is simple if the entity or construction is visible; for example, eliciting names for body parts, or parts of a traditional house is common in lexical elicitation. Likewise, the fieldworker can ask for words that are used to describe a collection of similar items; for example, a collection of birds is known as a *flock*.

Useful discussions of lexical semantics can be found in Saeed (2009:53–86), Lyons (1977), and Cruse (1986, 2004).

13.2.2 Sentence Meaning

Semantics also involves sentential meaning and interpretation, particularly the conditions under which a speaker can know if a proposition is true. In the philosophical or logical approach, the truth of a proposition is evaluated on the basis of the relationship between a pair of sentences, without taking into consideration the context or the speakers. For instance, Sentence A is said to be related to Sentence B because: A is the paraphrase of B; A entails B; A contradicts B; A is a tautology of B; or A presupposes B. This approach to sentence meaning – which treats sentences as "external objects" (Saeed 2009: 103) – is not immediately useful in descriptive fieldwork. As pointed out by Hellwig (2006:322), many of the tests for evaluating if and how a sentence A is a paraphrase of B, or A entails B, for example, are based on native speaker intuitions. It is difficult to judge the reliability of reported intuitions without already knowing the structure of a language quite well. Also, it is very unlikely that a natural text corpus will include examples of propositions opposed in paraphrase or entailment pairs or other semantic relationships.[2]

What is more useful for descriptive linguistic fieldwork is an approach to sentence meaning which takes into consideration the participants in the communicative act, as well as the implications of a sentence. This approach helps to identify why some sentences are acceptable while others are not. It is based on a number of factors: the meaning of morphemes; the connotations and denotations of an utterance under specific conditions; information packaging strategies based on clarity, speaker intention, and the speaker's assessment of the interlocutors' knowledge; and information backgrounding or foregrounding needs. For example, cleft constructions or pseudo-cleft constructions might only be accepted by speakers as felicitous when the clefted NP is known to both speaker and hearer, is part of their shared experience, or is in the immediate discourse, See Saeed (2009:105) and Matthewson (2004) for further discussion and examples.

13.3 Pragmatics and Fieldwork

This section presents aspects of conversational interaction that show how an utterance is "anchored" to context (Levinson 1983:55). The more varied the interlocutors and situations a fieldworker is able to document, the greater the likelihood of identifying morphological and syntactic anchoring strategies – some of which, no doubt, will have never been seen before by the fieldworker. The standard reference on pragmatics is Levinson (1983). Useful summaries of topics discussed in Section 13.3 can also be found in Schiffrin (1994) and Saeed (2009). For the challenges of doing pragmatic study on an endangered language, see Grenoble (2006).

13.3.1 Deixis

Deixis refers to the pointing out of referents that may be abstract (e.g. time) or concrete (e.g. person). Deictic reference is linked to a speech context. For correct

[2] See Matthewson (2004:401-410) for a different point of view.

interpretation of deictic indicators, the pragmatic co-ordinates or reference points must be known (Levinson 1983:58).

Levinson (1983) discusses four types of deixis: person, place, time, and social. Person deixis relates to pronoun distribution and interpretation. The particular pronoun used is dependent on the participants and on how they are situated with respect to the speech event.

Place deixis refers to a grammatical expression of movement within, location in, or measurement of space. Demonstratives such as English *this* and *that* are obvious examples. Many languages have proximal and distal determiners which specify whether an action involves movement towards or away from the speaker, or even above, below, or to the side of the speaker.

Time deixis is again linked to the speech context, reflecting when the proposition is encoded by the speaker and when the proposition is received by the hearer. Time deixis may be indicated through adverbials such as *now* and *then*, or through tense marking.

Social deixis refers to the grammatical encoding of social hierarchical position. Honorifics are a typical example of how social status is reflected in grammar. Pronouns and forms of address – in summons, vocatives, and titles, for example – can also vary depending on social status.

Deixis is also discussed in Section "Grammatical Categories Characteristic of Noun Phrases".

13.3.2 Conversational Implicature

Participants in a conversation are able to draw more information from an utterance than what is literally encoded. For example, if I say, "I've been looking for Mr. Bose all morning," my interlocutor might well reply, "I think he called in sick." The response is appropriate because my interlocutor infers that I'm hoping he can tell me where Mr. Bose is. The fieldworker should be aware that interlocutor responses are always shaped by conversational implicature and inferencing. Thus conversational responses will provide a rich variety of morphological and syntactic encodings of speaker attitude, such as surprise, deference, and impatience. Elicited responses may not provide such variety. The fieldworker should also be aware that rules of inference and implicature will differ from culture to culture, and can only be understood by actively using the language and observing the language in use.

A good example of implicature is reported for Mongsen Ao (a Tibeto-Burman language of Nagaland, India) (Coupe 2007:159–160), where agentive case marking is optional in some contexts. The question "Who drank the tea?", when accompanied by agentive case marking, highlights a volitional activity, and this highlighting implies that something out of the ordinary has occurred. The answer to this question would also have a marked agent, and again the agentive marking would highlight agency and therefore imply socially marked behavior.

13.3.3 Presupposition

The inferencing discussed in Section 13.3.2 focused on expectations that speakers have about conversational interactions. Presupposition refers to inferencing that is determined by specific grammatical constructions or words called presupposition-triggers. Here is a representative list of such triggers based on Levinson (1983:181–183).[3]

- Factive verbs, for which the embedded proposition is always true regardless of whether it is positive or negative, e.g., *Manny* ***regrets*** *finishing/not finishing the job on time.*
- Definite descriptions, which indicate that a speaker assumes the hearer can identify an entity, e.g., *I handed* ***the*** *book over to her.*
- Implicative verbs, which presuppose that an action did/did not occur, e.g. *She* ***managed*** *to open the door.*
- Change of state verbs, which describe a change in a presupposed previously existing state, e.g., *Jamie* ***stopped*** *bossing her colleagues around.*
- Iteratives, which imply that an action has occurred before, e.g., *The Martians* ***returned*** *to earth.*
- Temporal clauses, which presuppose that an action occurred at a particular time, e.g., ***Before Redford came to town****, Newman was all the rage.*
- Cleft Sentences, where the clefted portion is presupposed to be true, e.g., ***What I've lost*** *is my patience.*
- Non-restrictive relative clauses, where the relative clause is presupposed to be true, e.g., *The next candidate,* ***who is from Lubbock****, is coming in tomorrow.*
- Counterfactual conditionals, for which the counterfactual clause is presupposed to be untrue, e.g., ***If I were a rich man****, I'd build a big tall house.*
- Alternative questions, where one of the propositions is assumed to be true, e.g., *He was* ***either*** *Hispanic* ***or*** *Asian.*
- Wh-questions, where we presuppose that the proposition is true, e.g. ***Who ate the cookies?***

In many cases, real world knowledge of the speech situation can cancel the presupposition. Compare the following examples, taken from Saeed (2009:108). In (8a), the temporal clause triggers the presupposition that the proposition of the clause is true – i.e., she finished her thesis. In (8b), real world knowledge supports the interpretation that the thesis was never finished because the author was no longer alive to finish it.

8a She cried before she finished her thesis.
8b She died before she finished her thesis.

[3] This list is taken from an undated, unpublished manuscript by Lauri Karttunen (n.d) called "Presuppositional Phenomena" written at the University of Texas, Austin.

Thus context is also important in determining which presuppositions will hold. The fieldworker can use this information to understand why certain words and constructions appear in conversational interaction. If they are presupposition triggers, speakers are providing background information to an utterance.

Another important strategy that interacts with presupposition is information packaging. Entities that the speaker and hearer can both identify are often packaged as topics. So topics are known information, and what carries the new information in the clause is the comment. Kroeger (2004:137) identifies three main types of topics, and provides the following English examples:

- Contrastive Topic: A constituent has been moved leftward, or is co-referential with a "gap" (see underlined missing constituent), e.g., *This ice cream I like___, but the stuff we had yesterday was awful.*
- Left-Dislocation: The topic is not assigned a grammatical relation. A resumptive pronoun in the main clause is co-referential with the topic, e.g., *My friend John, a snake bit him on the hand and he lost three fingers.*
- External Topic: An old topic is reintroduced, e.g., *As for John, a python swallowed his dog.*

The topic/comment opposition should not be confused with the focus/not in focus opposition. Focus is typically contrastive, and carries information that goes against the interlocutor's expectations. In English, focus is indicated through intonation or cleft or pseudocleft constructions. See Chafe (1976) and Section 11.2.3.1 for further discussion of the topic/comment and the focus/not in focus opposition.

Obviously, not all languages will package topics or focused elements in the way that English does. However, these pragmatic concepts have important implications for syntax, and thus the basic terminology and concepts must be mastered by the field analyst so that the effect of information packaging on word order and other syntactic factors is not missed. For a clear (and somewhat simplified) discussion of the interaction of topics in Mandarin and Japanese, and focused elements in Indonesian and Russian, see Kroeger (2004:142–161).

It is not easy to determine whether or not shared knowledge plays a role in the selection of a particular syntactic construction. One method is to create a vignette, preferably one where the fieldworker and consultant know the characters and the relevant social networks, and to check the appropriateness of a set of constructions in that context. In this case, using a previously-analyzed narrative or conversation as the basis for the vignette is helpful, because both the fieldworker and the consultant share the same story-world information, and most of the set-up for the question is already present in the target language. For example, a fieldworker can use the information in the vignette to check on how contrastive focus is indicated. Say a conversation revolves around a group of friends getting together for a meal: a contrastive focus construction could be elicited by asking how one would inform someone that X was the one bringing the food.

With context provided by an existing text or created situation, one could construct several sentences connected to the discourse situations, and then ask a speaker if each is felicitous or true in that discourse situation. The native speaker's judgment would therefore be about felicity and truth, not about grammaticality. See Matthewson (2004:401–404) for how to interpret native speaker responses to this task.

13.3.4 Speech Acts

Speech Act Theory studies the ways speakers "do things with words" (Austin 1962). Speakers use sentences not only to relay propositions that are true or false, but sometimes use utterances to effect some change in their world. The following types of speech act are discussed in Bach and Harnish (1979):

- Performatives, which consist of a performative verb that overtly states the action that is being performed, e.g. *I baptize your child Mary*
- Constatives, which include statements that affirm, allege, announce, answer, attribute, claim, classify, concur, confirm, conjecture, deny, disagree, disclose, dispute, identify, inform, insist, predict, rank, report, state, or stipulate
- Directives, which include advising, admonishing, asking, begging, dismissing, excusing, forbidding, instructing, ordering, permitting, requesting, requiring, suggesting, urging, and warning
- Commissives, which include agreements, guarantees, invitations, offerings, promises, swearing, and volunteering
- Acknowledgments, which include apologies, condolences, congratulations, greetings, thanks, and acceptance (e.g. acknowledging an apology)

Speech acts of the same type have similar syntax. For example, in English, performatives have the following characteristics: they are in the present tense; always have a first person subject; allow for the insertion of "hereby" (as in *I hereby baptize your child Jacques*); and are in the indicative active mood. Also, they cannot be negated. Finally, the performative verb is taken from a special set of verbs which connote actions such as hiring, firing, naming babies, pronouncing judgments, and marrying two people. Since syntax and particular speech acts have predictable connections, the language analyst must be able to classify constructions according to their speech act category.

Another important aspect of speech acts is how they intersect with politeness strategies. In English, a command can indeed be issued through an imperative, e.g. *Quit jumping on the couch!* However, it is more common to find speakers using strategies that mitigate the force of a command by issuing it indirectly, that is, through the use of syntax that is not primarily meant for commands, e.g. the declarative, *You're going to fall off and break something if you keep jumping on the couch.*

Some languages use indirect speech acts for mitigation, but others soften the force of a speech act through changes in intonation or by selecting a softer imperative form that implores rather than commands. See for example, Coupe (2007:403) for Ao Naga, a Tibeto-Burman language of India. Other methods of mitigation may be the use of honorifics, pronouns which show deference, specific words like English *please*, and titles or terms of endearment.

As discussed in Burt (2008), in the case of language shift and language contact, researchers may find politeness particles used differently by speakers who are situated at different points along the language shift continuum.

13.3.5 Conversation Structure

When studying conversation, both structure and pragmatics should be considered. Although a lot more comparative work is needed, it does appear that conversations universally have a predictable structure. In the simplest instance, two speakers, A and B, will take turns talking. Each turn in the conversation, called a turn constructional unit (TCU), may be a complete clause or set of clauses, a phrase, or even just a single word. The end of a turn constructional unit is called the transitional relevance place (TRP), and this is where the bulk of the intersection of conversational structure and syntax lies, as seen in the following list:

- Overlap: When speakers overlap in talk, it appears that the overlap does not occur at random syntactic points, but rather, at constituent boundaries or at predictable points within a constituent. For example, overlaps of relative clauses occur at the head of that clause. See Fox (1987) for further discussion.
- Turn Constructional Unit extenders: It is possible to extend one's turn by using conjunctions and other connectors which allow for possible additions, e.g. *that's where I went and …*
- Selection of next speaker or Floor return particle: To end one's own turn and force another speaker's turn, the speaker may use a tag question (e.g. *right?*) or a floor return particle (e.g., *hmmm*) which prompts a response.
- Repair system: To correct one's own or the interlocutor's production, a variety of repair devices can be used. Echo questions can be used to clarify or correct an interlocutor's statement, e.g., *You went to Spain?* An explication question may also be used e.g., *You mean you went to Spain?* Using right dislocation, the speaker might also clarify pronominal reference, e.g. *He's coming tomorrow, John is*. For an excellent typology of conversational repairs see also Geluykens (1994).
- Adjacency pairs: Adjacent turns often have predictable syntax. Thus in most cases declaratives, interrogatives, and imperatives will all be responded to by declaratives. Since the responding item in an adjacency pair often repeats a part of the previous turn, it can be used to determine syntactic constituency.
- Pre-requests: When a turn threatens the negative face of the interlocutor, a speaker might mitigate that threat through the use of polite past modals, e.g. *I was wondering …* or *Could you …*
- Controlling the flow of topics: In order to revive a topic previously under discussion, speakers may employ left dislocation, e.g. *John, I like him.* An adverbial like *Anyway …* can be used to end the current topic and either introduce a new

topic or bring back an old one. Other strategies for introducing new topics include the use of interjections, e.g. *Listen! Did you hear about John?*; adverbial phrases, *e.g. By the way…*; or existential clauses with the frame *There+BE+ITEM*, e.g. *There's this guy...* See Schiffrin (1994) for further discussion.

We have used English data to illustrate that the syntax between turn constructional units is to some extent predictable. In other languages the work of conversational repair or topic flow may be carried out through morphological means. For example, see Mirzayan (2008) is a study of self-repair in the polysynthetic language Wichita, a Caddoan language of Oklahoma. In either case, because of the abundance of contextual information, conversations are of the utmost use in determining the meaning of constituents. This is especially true of those lexical or morphological constituents that are abstract or have multiple functions, e.g., English *anyway*.

13.4 Texts

While the collection and analysis of texts – that is, naturally occurring discourse – cannot fully replace elicitation as a fieldwork technique, the data from texts are fundamental for an understanding of language structure, as reflected in the words of these seasoned fieldworkers[4]:

> Texts are the lifeblood of linguistic fieldwork. The only way to understand the grammatical structure of a language is to analyze recorded texts in that language (not by asking how to translate sentences from the lingua franca). (Dixon 2007:11)

> An account of any language needs to be based primarily on a substantial corpus of continuous spontaneous speech. (Crowley 2007:120)

> The text collection seeks to show the language as it really is, and among other things provides a corpus against which the grammar's claims can be tested, and which subsequent linguists may scrutinize for generalization overlooked by the original grammarian. (Evans and Dench 2006:12)

> There is general agreement, I believe, that a grammar should describe a language as it is spoken… Thus, in fieldwork, the need for working with spoken language of a variety of genres has long been recognized; grammars that do not draw richly from such material are probably unlikely to attain the goal of describing the genius of a language [referring here to Sapir (1921)]… These are, I believe, absolutes… (Rice 2006:23)

Given these views, it is surprising that, at the time of this writing, linguistics places a limited value on text collection. Mosel (2006:52–53) attributes this to three factors:

- Many linguistics departments do not recognize descriptive linguistic fieldwork and its resulting products as valid Ph.D. thesis material.
- The field of linguistic typology typically uses single sentences for the cross-linguistic study of grammatical phenomena, and therefore data is seen as more

[4] See Section 13.4.2 for a list of naturally-occurring speech that qualifies as 'text'.

efficiently derived from targeted questionnaires rather than from texts. Typologists require vast amounts of comparable data for cross-linguistic study, and it is too time consuming to mine texts for these data.

As we will show in this section, even though the prospect of transcribing, translating, and annotating texts seems daunting, there is no getting away from the fact that data from texts are a necessary product of fieldwork. We also provide a methodology for text collection and analysis, pointing out some problems and pitfalls in the use of textual data.

13.4.1 Advantages of Text Collection

There are many reasons for creating, providing access to, and using text collections in language analysis. First, texts – especially narratives and procedural texts about traditional activities – may become the sole record of the oral tradition of a community (Crowley 2007:128). Thus native speaker communities can recognize the importance of text documentation for the maintenance of their language and culture, and may encourage this linguistic activity. As discussed in Yamamoto (1998:232) and in Chapter 6 of this book, the type of data collected should be determined by what the community thinks is important. Providing recordings of traditional oration in accessible formats is one way that the fieldworker can fulfill his or her commitment to community needs.

Data from naturally occurring speech is reliable in that they have not been corrupted by priming or by other translation or elicitation effects, since speakers concentrate on the stories rather than on the constructions they are producing. Furthermore, some of the linguistic structures found in texts may never emerge in elicitation; for example, epistemic modals might occur regularly in elicitation, but miratives might surface only in conversation. Another good example of this, pointed out by Amha and Dimmendaal (2006:431–433), is that texts must be studied to understand the distribution and meaning of converbs in Nilo-Saharan and Afroasiatic languages, because speakers use converbs in response to specific types of conversational moves. See Chelliah (2001) and Section 12.2.2.8 for other examples of the need for textual data.

While data from texts do not include negative data (that is, structures which are *not* grammatical in the language), they are a useful springboard for further analysis and can fruitfully be used along with elicitation. We cite an extended example from Crowley (2007:128) to illustrate this point:

> Of course, any new constructions which appear in textual data can be supplemented by additional elicitation. Elicited translations from English or some other language may point to the existence of a separate category of past tense of verbs in a language. With careful elicitation, you can complete the full paradigms for the past tense and you may think that you have done all there is to do. However, once you start recording stories, you may find that past tense meanings are occasionally expressed by quite different forms. Further investigation may reveal that the original past tense paradigm that you recorded only related to the immediate past, and there is, in fact, a completely separate paradigm for the distant past. However your textual attestations of this new paradigm will possibly not provide you

with forms for the full paradigm. You would therefore need to supplement the data from your text with further elicited data in order to complete this new paradigm.

Healey (1975:355) suggests that using the text as a guide to elicitation is helpful in monolingual situations. He recommends that, when an unusual morpheme or construction is seen in a text, the fieldworker should elicit five other constructions of the same type for each morpheme, word, and clause that is new. Working in this way, Healey says, the speaker will sense that the fieldworker is progressing in a systematic fashion.

A text collection serves as an evidence bank against which a linguist's claims can be independently verified Mosel (2006:53). Furthermore, the text corpus can be used to improve on previous descriptions and analyzes. Mosel explains how she failed to describe the Tolai (Austronesian of Papua New Guinea) particle *iat*, but that this particle was identified and described in later work by another researcher, an advance that was made possible by the existence of a text corpus.

Texts are also useful in determining whether or not a set of speakers is using the same dialect or language. To begin, record a personal narrative from speaker A. Speaker B can then listen to this narrative and answer questions about it. Scoring the responses for lexical and grammatical similarity with the variety in the narrative will give the analyst some idea of the distance between the varieties used by Speakers A and B. This use of narratives is explained in Grimes (1995:18). When samples of naturally occurring discourse are collected from speakers of both genders and a variety of ages and genres, it is possible to document the effects of age grading or register modification.

Recording narratives can help break up the monotony of fieldwork sessions. It puts the speaker in control of the session, so that even further discussion of the text for purposes of transcription and translation are speaker-centered activities. In our experience, consultants enjoy working with natural discourse produced by other speakers. They are curious about what was said and how it was said. In fact, native speakers tend to accept and process utterances from other speakers much more readily when those utterances are taken from texts. It is important to use data in linguistic descriptions that is acceptable to speakers if we want those linguistic descriptions to do some good for the community. Ameka (2006:73–74) describes a situation in which, when speakers carefully examined the sentences used in articles on their language, they considered many of the examples to be ungrammatical. Even if we assume that such examples were collected using careful elicitation methods and judicious use of grammaticality judgments, and even if we state that speakers can never produce ungrammatical constructions in their native language, or if we claim that they speak different idiolects and therefore have differing judgments about the same construction, we still leave the native speaker in doubt about the validity of linguistic research when that native speaker finds the majority of cited language examples to be unacceptable. In Ameka's (2006:74) words, "It makes one wonder sometimes about the empirical bases of some theoretical claims".

Finally, collecting texts is important because sometimes that is the only data that speakers can produce. In some endangered language situations, younger speakers may not know traditional stories and ritual language, but they may nonetheless be able to carry on simple conversations. But the reverse can also be true: younger speakers may have memorized texts from elders, but may not be fluent enough to converse in the language.

13.4.2 Types of Texts

Foley (2002:136) defines a text as, "a body of language behavior generated continuously over a period by the informant and recognized as an integrated whole." This is a helpful definition for the fieldworker, because it includes virtually any connected naturally occurring utterance, and validates what fieldworkers know from experience: even the most informal interaction can productively be mined for grammatical information. An extensive list of genres, with subtypes, can be found in Dwyer and Mosel (2001). Here is a representative list of traditional texts a linguist might collect:

- Creation stories or myths
- Folk stories or fairy tales
- Genealogies
- Legends
- Parables, sayings, proverbs, riddles, and jokes
- Ritual ceremonial texts or prayers
- Procedural texts such as: how to cook X, how to build or make X, how to catch X…
- Songs
- Poems
- Plays

Here is a representative list of non-traditional texts:

- Anecdotes
- Life experiences (see the list in Crowley 2007:126)
- Biographies or autobiographies
- Stories about professional activities
- Descriptions of pictures or video-clips
- News broadcasts from radio or TV
- Tapes with messages
- Letters, good wishes (see Dorian 2001:140)
- Re-tellings of stories
- Religious or moralizing sermons
- Conversations
- Any kind of non-traditional literature, songs, poems, plays, sayings, proverbs, riddles, or jokes

Other examples of spoken texts can be found in Kibrik (1977:61), Payne (1997:356), and Crowley (2007: 126).

Monologues are problematic in terms of naturalness. Roman Jakobson (Hoijer 1958:590) remarked that most speakers are not accustomed to the "monologue" speech activity, and that it is even more artificial than thinking and talking about the target language.

While written texts cannot take the place of spoken texts in a fieldwork corpus, they should not be ignored. They are often non-traditional and therefore circumvent issues that occur with sensitive cultural information found in traditional texts (Mosel 2004). Written texts are useful for understanding prescriptive grammatical rules, adding to word lists, providing material for discussion with consultants, and revealing the grammatical structures used in more formal communication. Furthermore, encouraging literacy in native cultures can aid in language maintenance, while at the same time creating new textual data. Francis and Gómez (2009), for example, report on promoting creative writing as part of a language maintenance project; this resulted in an informal Nahuatl bulletin composed of short essays on Nahuatl culture. These essays enriched the type of discourse data available to the researchers.

Songs, poems, and plays should also be used with the understanding that a literary form may influence the linguistic structures used. This is true in Manipuri, where archaic language and stylized intonation are typical of literary genres and dramatic delivery. That is why it is best to analyze these at a later stage of collection, when the spoken language is understood.

Silko (2001:161) points out that it is not always easy to categorize texts correctly based on one's own limited cultural experiences. For example, a set of poems might have religious significance for speakers, but might be appreciated only for their aesthetic value by the fieldworker. Distinctions between gossip, religious texts, and historical texts might be blurred in some communities. Similarly, the line between scientific and cultural texts may be fuzzy for some speech communities (see Albert 1972). So if speakers ask fieldworkers not to document religious texts, the fieldworker must be clear about which texts fall into that category.

Silko (2001) also points to further potentially faulty assumptions. It is usually thought that native speakers do not mind having their words written down and then analyzed; as discussed in Section 6.3, this is not always the case. Also, it is commonly thought to be good methodology to record information about when a story should be told, or when it might not be appropriate to tell a story; but some speakers might question this practice, because to them stories are always relevant. Silko says that, in the Laguna language (Keresan of New Mexico), it is said that story-telling "goes on constantly" and is a way of connecting present experience with past experience.

In terms of the quantity of texts to collect, there is a conflict between what is good for analysis and what might be of prime importance to speakers. As noted above, a collection of traditional narratives is a precious community resource. The researcher, on the other hand, might find the same syntactic patterns repeating themselves in such narratives. For example, narratives tend to utilize the past tense and/or perfective aspect. Thus other types of texts, such as conversations, will be needed to flesh out the study of tense and aspect. As discussed in Crowley (2007:129), collecting narratives from a variety of speakers of all ages will advance both agendas, since individual variation will add interest to the syntactic and morphological data collected. Some speakers may use more evaluative clauses in their narrative, for instance, and thus provide morphology that reflects subjective reactions to a situation. Collecting a variety of texts from a variety of genres will also aid in language revitalization efforts, much more so than just collecting traditional narratives (Amery 2009).

13.4.3 Obtaining Texts

Obtaining a narrative from a speaker can be either a controlled activity or a free activity. A controlled activity is one where the fieldworker can predict something about the lexical or grammatical content of the resulting text, because the prompt has been carefully prepared to elicit a specific type of text. Speakers may be asked to:

- Look at a book and tell the story. A common prompt used for this is "Frog, Where are you?" by Mayer (1969), a 32-page wordless picture book. Since the book was designed for 3–6 year olds, it may not be appropriate for use with all consultants, but it is thought to be acceptable for use in most cultures.
- Use stimulus prompts, especially picture prompts (see also Section 12.2.2.2). Sutton and Walsh (1979:6) report that when there are so few fluent speakers of a language remaining that it becomes difficult to elicit connected clauses, then culturally appropriate pictures can be used as prompts to jog speakers' memories. For all speakers, it is easier to talk about something concrete – an object or a picture – rather than something abstract. Rather than ask for "stories about your family", the fieldworker could use consultant family pictures and ask questions about those (Jackson 1987:98).
- Watch a movie and narrate the events viewed. A common prompt used for this is the "Pear Story", a 6-min silent film created by Wallace Chafe and his colleagues at the University of California, Berkeley in the 1970s (Chafe 1980). The predictability of plot and repeated occurrence of the same entities and actions in the film make translation relatively easy. This task allows the fieldworker to quickly collect samples to compare speaker varieties and find alternate ways that speakers can talk about the same scenes. The Pear Story can be downloaded from http://www.pearstories.org/. Since linguists often use the Pear Story as a prompt, it would be useful to find out if translated Pear Stories already exist for languages related to the target language. However, speakers may be puzzled or uninterested in this culturally foreign story, especially if undue interest is given to it at the expense of traditional narratives.
- Translate a written story from the contact language into the target language.
- Provide a version of a well-known story.
- Read or paraphrase a written story: The advantage of this prompt is that, in the free translation, the sentences will be complete, and thus easier to parse and gloss – unlike unguided natural speech, which is less predictable. The disadvantage is the unnatural setting which may result in unnatural or prescriptive forms.

One interesting follow-up activity is to get a spontaneous retelling of the same story at a later date in order to compare the two versions.

Free narrative tasks include telling a personal narrative or a traditional folk tale, or producing a monologue. The personal narrative is easy to collect. It can be elicited by asking a speaker to talk about something exciting that has happened to them. Labov's (1972:93) request to speakers to talk about a near-death experience can be modified with excellent results. The fieldworker can watch and listen for recent events in the life of the consultants and ask about those. In one situation, I [Chelliah] was interviewing a very shy monolingual speaker who had just taken her first trip on an airplane. I heard from others that she was planning a return journey by train, and that she would never take a plane again. I asked her to tell the story of her flight, which she did with great animation. Another prompt I have used to get a personal narrative is, "What's it like to work at ___(fill in the relevant workplace here). Has anything exciting happened to you at work?" For comparable texts from different speakers, one good approach is to ask about their families. Speakers will tell you only what they want you to know, so there is no danger of embarrassment, or of forcing information from them. Of course, no sensitive topics should be pursued.

The prompts for free narrative elicitation must be culturally appropriate and effective. Milroy (1987:40–41) reports that asking for descriptions of a near-death experience from "world-weary" speakers in Ulster (Ireland) elicits a straight factual account, but not the high energy, fully animated response that Labov intended to elicit, perhaps because speakers are so constantly faced with near-death experiences. In all cases, it is important to let speakers know that their perspective, their stories, and their narrative style are of interest to the fieldworker. This will encourage speech even in those cases where both the speaker and fieldworker already know the answer to the posed question. Another point is that if the speakers don't know what is going to be done with the conversation or stories, they may be less inclined to talk to the fieldworker. Jackson (1987:98) relates the story of a man whose mother was a great teller of family stories. When he tried to elicit these from her using family photographs as prompts, she dried up. The reason given was that she was not told why her relative – who had heard these stories all his life – suddenly wanted to record her stories. Was he perhaps testing her to see if she was senile?

With free narrative collecting, it is important to get as natural a response from speakers as possible. One issue here is that since the speaker knows that the fieldworker does not know the language well, he or she will not perform in the same way as when there is "immediate intelligibility" (Scollon 1979:10). It is thus important for the fieldworker and other assistants to look interested in the narration; if possible, it is helpful to have other native speakers of the language present.

If the fieldworker is interested in rhetorical analysis, it is important to select the right speaker to provide a monologue – say, someone who is accustomed to public speaking.

Most free narratives are not planned or elicited through prompts. If there is a thriving speech community, recordable stories will pop up at unexpected moments.

The fieldworker should be ready at all times for these "incidental contributors" with a recorder, gift, camera, notebook, and release form (see Section 8.5) ready to go.

In endangered language situations, speakers may not recall enough of the target language to produce connected discourse. Because speakers do not hear the language on a regular basis, they may need some quiet time just to think in the target language and prepare for the task. Another method might be to re-elicit texts that have already been collected.

As Foley (2002:135) and many others have asserted, conversations must be included in a descriptive corpus because it is in such texts that social distance and speaker intention surface, and these are often expressed through morphological and syntactic features not found in other texts. Just as in the case of narrative elicitation, the elicitation of conversations can be controlled or free. Examples of controlled conversations are:

- Games or activities which force question-and-answer exchanges, such as "twenty questions", interviews, and the like. Other examples of ways to encourage conversation between speakers in a guided way can be found in any Teaching English as a Second Language activities book.
- Scripted conversations. These are interactions between two or more play-acting speakers, using scripts. The scripts for these interactions may be:

  (a) A transcript of what the speakers previously said extemporaneously
  (b) A script created by the speakers
  (c) A script created by the fieldworker, but checked by speakers for accuracy
  (d) A script created for one speaker, with the other speaker answering extemporaneously

Of the script types listed here, most are appropriate only for the most creative and outgoing speakers. The more guided scripts are useful for endangered language situations (see Caldecott and Koch n.d.). The fieldworker needs exactly the right combination of speakers for recorded conversations to approximate natural conversation, but if s/he has two or more speakers together, it is certainly worth a try. If it works, one useful goal is to produce a series of "conversations" as part of a conversational manual for the fieldworker's own use, or for the community.

Eliciting natural-sounding conversations is difficult. It is strange to engage in conversation on demand, so only a particular type of speaker will be able to "perform" under fieldwork session conditions. Some speakers can engage in a fairly realistic conversation, but in many cases speakers will produce what amounts to alternating monologues. The more unnatural the data is, the less useful it will be. For example, a very stilted conversation may use simplified syntax in order to help the fieldworker understand what is being said. Here are some methods discussed in the sociolinguistic literature that help with obtaining natural conversations:

- Peer Group recordings: This is a technique employed by sociolinguists to record natural interactions. The idea is to find a group of speakers who self-select to be together. The interviewer is a listener; the content of the conversations comes primarily from the group members. See Cukor-Avila (2006) and the references cited there for discussion of the data resulting from such recordings. Jackson (1987:94) describes how listening can be used in this kind of recording, "If you keep your mouth shut whenever possible, if you listen rather than lecture, if you don't load the conversation, if you follow their lead, you get taken places you didn't know were there. If the places are dead ends or boring or irrelevant, you can always steer the conversation back to where you hoped it would be going, and you can do it directly [like this]: But before we talk more about the truck tires, I'd like to hear a little more about the time you were on the ice floe. Just how did you get off of it…?"

- Ethnographic research, including participant observation: In this case the fieldworker has permission to live with a community and take part in day-to-day activities. The fieldworker speaks the target language to some extent.
- Telephone conversations: These can be recorded, but the fieldworker must get prior permission from both interlocutors so the transcript can be used for publication. Also, note that the interactional routines of phone conversations are not the same as face-to-face conversations.
- Interviews: The interview can be a directed question-and-answer session, or a less formal and more conversation-like speech event. In the less formal interview, the interviewer's voice should be heard as little as possible. For more on interviewing techniques see Jackson (1987:79–102), where he discusses topics such as interviewer register shifting; facial expression; effects of turning the recorder off and on during the interview (something he discourages, because it reveals what the fieldworker thinks is important); follow-up questions; and the art of "acting natural".
- Community recordings: The fieldworker trains an interested native speaker in recording and basic cataloging methods. The native speaker then takes recording equipment to a community site of his or her choosing and records interactions at that site. The native speaker notes all pertinent ethnographic and demographic information for the project. This method is useful in areas that cannot be accessed by the fieldworker for reasons such as political unrest or lack of law and order.
- Fieldwork session elicitation: It is possible to record conversations between participants at field sessions. However, the speakers must know each other well or be curious enough about each other for the conversation to go from initially stilted to more natural.

Sometimes it is useful to elicit narratives when there are several speakers at the fieldwork session. I [de Reuse] remember one speaker who volunteered a "pack-rat recipe" in Apache. That sounded quite exciting to me. However, on listening to the text with another speaker, it turned out to be a somewhat rambling and repetitive statement about the fact that Apaches used to eat pack-rat. Having several speakers present at the recording would have helped control the rambling nature of the performance.

13.4.4 Recording and Analyzing Narratives

Text collection or discussion of texts previously collected should be a regular part of most fieldwork sessions. The object is to produce a richly annotated collection of texts that can be mined for data and used for dissemination of cultural and linguistic information. We review seven steps for deriving usable data from texts: recording, rough translation, transcribing, word-for-word translation, constituent analysis, free translation, and morpheme analysis. The order in which we present these steps may deviate from the order followed by other researchers. For example, Scollon (1979) gives the following as an example of the process he followed for work on Chipewyan (now called Dëne Sų̀łiné, an Athabascan language of the Canadian North):

- STEP 1: Record the speech event in the target language, and get a free translation in the contact language, if possible.
- STEP 2: Right after recording, make notes about the recording, explaining ethnographic details (e.g., setting) and paralinguistic factors (e.g., gestures).
- STEP 3: Transcribe the recording without the help of a native speaker. The resulting "irregular" transcript will have half phonemic and half phonetic transcription.
- STEP 4: Gloss as much as possible. If there is some English in the text, transcribe that part.
- STEP 5: After letting that transcription "stew" for a while, complete the narrow phonetic transcription. The fieldworker is aided by the fact that, by this time, he or she has heard the speaker on several occasions and has heard the text again.
- STEP 6: Write up the text in phonemic transcription with detailed discourse transcription (breath groups, terminal intonation) and complete glosses.

A different way to structure the annotation process is suggested in the discussion below. See also Lehmann (1982) and Bow et al. (2003) for more suggestions on annotation.

13.4.4.1 Recording

Before electronic recording equipment was easily available, texts were dictated to fieldworkers whose expert transcriptions skills allowed them to faithfully record them. It is questionable, however, whether a dictated text can be produced naturally, since speakers tend to change their pronunciation when speaking at a slow careful rate; this makes dictated data of limited use for detailed phonological study (Boas 1917:1–2). Today, with recording equipment so readily available, dictation as a method of data collection is unthinkable. Recordings allow transcriptions to be completed with native speaker input; transcriptions can be re-checked if necessary, and the recordings can be analyzed acoustically to identify intonation patterns.

In addition to the record-keeping for recordings discussed in Section 8.6.1, note should be made of the audience and other conditions under which the narration or conversation occurs. These "situated performances" cannot be understood without information on where and to whom they were said (Scollon 1979:3).

Many fieldworkers have horror stories about recording well-delivered narratives or natural conversations and having the recorder stop recording because there were no batteries in the microphone (Crowley 2007), recording over another narrative, or running out of room on the memory card. Needless to say, equipment must be checked before each field session.

Other problems with recording naturally-occurring speech are that speakers may not be able to tell a coherent story if they are drunk, forgetful due to old age, or just bad storytellers. It is still a good idea to record such speakers, if no better speakers are available. It must be kept in mind, however, that it is very difficult to translate a story if it is abbreviated, if the episodes are out of sequence, or if two or more stories are mixed together. The recording device itself does not usually distract or bother the speaker, especially after a few minutes of recording. At first, though, speakers may use a more formal register, as if they are aware of a change in their "audience" (Tedlock 1983:292; Jackson 1987:87–89).

When recording narratives and conversations, it is best not to interrupt the speaker with questions; the speaker may switch to a more formal register or perhaps stop narrating all together. Of course, judicious use of back-channeling cues by the fieldworker is always useful, especially when modeled on how other members of the community "listen" to a narration.

13.4.4.2 Video Recordings

Video recordings add a new dimension to data collection. There are pros and cons to the use of video. Fieldworkers who caution against the use video include Dixon (2010:318), to whom the camera "gravely disturbs the chance of establishing a close relationship between the linguist and speech community". Fieldworkers report that while speakers are not intimidated by tape recorders, the video camera makes them self-conscious. Many speakers want to rehearse before speaking in front of the camera. They may prefer to read from a script, and they may well want to dress up for the video camera.

There are special challenges with getting permission for video recording. One reason is that researchers or film-makers might make a film but fail to get permission for dissemination, or may not give due credit or pay royalties from screenings. Also, in some communities, the performance of certain cultural or religious events in front of outsiders may be prohibited, according to tradition. See E-MELD (2006b) on the permissions needed for video recording.

Finally, cultural mores may restrict women from using video recorders. Sadaf Munshi (p.c.) tells us that even though she had permission from a bride and groom and their immediate families to record their wedding ceremony, and even though there was a commercial video recording being made at the same time, an older female member of the community nonetheless prevented Munshi from recording. This elder's objections were based on Munshi being a woman herself, and also on the presence of other women who would be recorded.

Austin (2006:91) and Nathan (2007:3) list these additional problems with video recording: the cost, expertise, and human resources needed for creating and processing video before it can be used for analysis; the prohibitive digital space required for transferring and storing uncompressed video files; and the lack of guidelines on inclusion of video in archives.

On the other hand, video recording can be a useful tool for analyzing interactional data; as discussed in McConvell (2003, 2007) and Wittenburg (2007), with video one can identify the participants in conversations; document the speech event setting; record paralinguistic features including gesture, facial expressions, and body posture; and record lip rounding or spread. All of these can be helpful in analysis, e.g. checking on phonetic detail, or examining the pragmatic import of an interaction. Of course, video is recognized as a central tool in the documentation of language and culture, and it is useful in revitalization efforts.[5] Since speakers tend to value video recordings over audio or written text, in some cases it will be essential to record videos in order to satisfy community priorities (Wittenburg 2007:4). How much video recording one does should correspond to how endangered a language is. If a language is highly endangered, then Wittenburg's (2007:5) call to, "make as many video recordings as possible to document as much as we can before it is too late," makes sense.

It takes some skill to learn video camera placement, so this should be practiced before field recording takes place. The most common error reported by first-time video makers is using the internal microphone on the video camera; an external microphone should always be used with the video camera. It is also highly recommended that a digital recorder be used for backup audio recording. An excellent guide on camera placement, lighting, synchronizing the beginning of video and audio recording, and other basics of video use for language documentation is Cholin (2004). A good source on lighting and filming movement is Jackson (1987).

13.4.4.3 Rough Translation

Dixon's (2010:322) advice is to never try to record a text and then translate it later without native speaker input. To avoid having to do this, the fieldworker might record speakers' summaries of what they've just said or talked about at the time of recording. This meta-text will help with giving the narrative a working title and provide a scaffold for the actual translation task, as it provides a preview of the lexical items that will show up in the text. Remember that the speakers' translations can only be as good as their proficiency in the contact language. For example, if they

[5] For example see Tim Montler's Klallam page, where he presents annotated videos of Klallam elders interacting: http://www.lingtechcomm.unt.edu//~montler/Klallam/videos/index.htm.

don't know the word for 'pear' in the contact language[6] – perhaps they've never seen a pear before – they might use something like *mango* instead. Similarly, the tenses and aspects that are present in the contact language cannot always be taken as accurate translations of the target language. Several of my [Chelliah's] Manipuri speakers use *intend* for the future, so, *He intends to go* means 'He will go.'

13.4.4.4 Transcribing

To transcribe a text, the fieldworker typically sits with a consultant, plays a recording, and has the speaker slowly repeat what is on the recording so that he or she can write the utterance down using the IPA or a practical orthography. In the first few weeks, the activity can be exciting, because phonetic and phonological rules are being revealed to the fieldworker and these revelations can be discussed with interested speakers. Foley (2002:136) gives the following steps for this procedure, to which we have added a few suggestions:

- STEP 1: The fieldworker plays back no more than 10 s of the text. It helps to stop at a pause in the speech signal, as this often corresponds to the boundary of a syntactic constituent. Software such as TRANSCRIBER or ELAN is useful for this purpose. With TRANSCRIBER, the sound signal can be broken into breath groups easily, and each group can be played back with a mouse click. If desired by the fieldworker, the transcription can be written directly into the program. (There are some problems to watch for with this: diacritics and phonetic characters cannot be (quickly) keyed in; and there is the obvious danger of not saving or of losing typed work due to some kind of oversight or electronic glitch.)
- STEP 2: The fieldworker asks the native speaker to repeat the played portion. If the speaker on the recording is different from the person helping with the transcription, it should be ascertained whether or not the recorded material is potentially insulting or taboo. Healey (1975) recommends erasing offensive materials altogether, and of course, some things – such as defamatory material about an individual – should absolutely be erased. But not everything that is potentially offensive to the fieldworker will be offensive to speakers. In our experience, speakers differ widely on what shocks them. Also, native speakers can themselves filter material in the way they deem appropriate by either ignoring the material that they do not want to translate or by asking the fieldworker to skip to a different portion of the story. Some speakers are hypersensitive and see sexual innuendo or insult in every sentence. The fieldworker must adjust to different consultant personalities.
- STEP 3: The fieldworker needs to gauge if the transcription assistant is able to repeat exactly what is on the recording. Not everyone can do this. Some assistants will provide "corrected" versions of the pronunciation and grammar of the speaker on the recording. Others simply do not understand the task and may look to the

[6] Needless to say, an important word in a "Pear Story" retelling.

fieldworker for help. Others are fantastic at this task and will repeat the exact same string several times without tiring. The consultant should be a good fit for this task.
- STEP 4: The fieldworker repeats what the native speaker has just pronounced.
- STEP 5: The fieldworker transcribes the utterance if the native speaker agrees with his or her pronunciation, or asks for another repetition. It may be necessary to listen to the recording again.

Healey (1975) estimates that each hour of speech can take up to 70 hours to transcribe and roughly translate. This time-consuming and tedious task must be mixed in with other tasks in order to maintain speaker interest.

Some texts are easier to transcribe than others. Procedural texts are among the easiest because of the inherent repetitions. In my [de Reuse's] experience, traditional or folktale narratives are somewhat harder, biographical narratives still harder, and historical narratives very hard. Conversations are hard because of interlocutor overlaps, but easier if the turns are short.

In any case, the first text is always the hardest to transcribe. After the first few texts, text transcription will be easier, but it will always be a challenging task for both fieldworker and speakers.

If possible, after the fieldworker has spent time transcribing texts in the traditional way, he or she should investigate other ways of transcribing. Here are a few suggestions:

- A native speaker who knows the practical orthography could be hired to transcribe texts. The fieldworker will need to provide the right equipment for this. It might be useful to take an inexpensive laptop to the field for the consultant to use for this purpose, since transcriptions could then be directly entered into software such as TRANSCRIBER.
- If time is short, the fieldworker might record a speaker repeating a text at slow speed. This recording could be used to later transcribe the text (Paul Kroeger p.c.). One recorder will be necessary to play back the recording, and another to record the slow speech repetition.
- The fieldworker could try to transcribe the text on his or her own, and then check the transcription with a native speaker. Only the simplest texts can be transcribed with no native speaker input, and this should only be attempted after much practice and experience with the target language.

As discussed in Himmelmann (2006), the fieldworker will have to make some basic decisions when working with utterances that are longer than a single word. The first concern is determining a useful definition of a word boundary for the target language. The speaker may have some intuitions about what constitutes a word, but this may be determined on orthographic rather than structural considerations. We have worked with speakers of Tibeto-Burman languages who have been taught to write their language, so that the words are short following the argument that long words are hard to read and are aesthetically unsightly in printed form. Furthermore, orthographic conventions themselves may be inconsistent, e.g. the English compounds *blackbird* and *black fly*. See Himmelmann (2006:255) for further discussion. The fieldworker should expect that conventions for transcribing compounds, clitics, and idiomatic collocations (e.g. English *kick the bucket* or phrasal verbs such as *carry on)* will evolve as fieldwork progresses.

13.4.4.5 Word-for-Word Translation

Getting a word-for-word translation for a text may require the assistance of a different consultant than the one who helped with transcription, because this task requires different analytical thinking and translation skills. Some speakers cannot repeat what another speaker has said. Some are very good at word-for-word translation, while others simply cannot do it.

It is very helpful to have a second or third contact language in common with the consultant: if they cannot translate into one language, they might be able to find an appropriate word in another. When the contact languages are similar, this can get confusing! I [Chelliah] have worked with speakers of Lamkang (Tibeto–Burman of India) who usually translate into English but often provide Manipuri or Hindi translations. The Hindi translations are easy for me to recognize; however Manipuri is close enough to Lamkang phonologically – at least, the Lamkang pronunciation of Manipuri makes it so – for me to mistake the Manipuri translation for more Lamkang text.

There are several ways to complete a word-for-word translation. Traditionally, the fieldworker sits with a native speaker and fills in the translations. Along the way, questions can be asked about various aspects of the text, from culture to grammar, so this is a fruitful activity for the fieldworker. It is also possible to have literate native speakers fill in the glosses; the fieldworker can then study this translation and later meet with the native speaker to ask questions. Because the translations do not need to be from the same person who provided the text, the fieldworker can get texts from monolingual speakers, which may be more authentic and thus more valuable.

13.4.4.6 Constituent analysis and Free Translation

A rich and useful text translation results from repeated study of a text, and input from the fieldworker's varied experiences with the text. Evans and Sasse (2007) list the following types of information that go into text translations:

- Knowledge of how the language works
- Information from gesture
- Comparison with others who have told the same story
- Other information from the discourse setting
- Remarks made by the narrator after the discourse event

Even if a consultant has a good grasp of the contact language and can provide a quick translation, this does not mean that the fieldworker can consider translation quick, easy and done. Rather, many strands of information should be used to enrich initial translations.

A fieldworker's early understanding of the target language, while helped by preparatory reading on related languages, is limited enough that constituent analysis on data from texts may be difficult. It is advisable, as suggested by Foley (2002:134), to first write a grammatical sketch of the target language which can then be used as a guide to attempt the analysis and translation of texts. Mosel (2006), a useful overview on how to write a grammatical sketch, points out that a sketch is a work in progress. The first sketch need not take long to write. It would note the most obvious phonological, morphological, and syntactic patterns found and would include questions for further investigation. The sketch would be revised as more information becomes available through text collection and elicitation on the basis of texts.

It is helpful to have a clause-by-clause translation before attempting a morpheme-by-morpheme analysis. For example, one can be on the lookout for past or future tense forms if it is known that the translation suggests past or future action. Matthewson (2004:348) makes the important point that the translation is a "hint of the analysis" and not the analysis itself.

To tap into speaker intuitions, the fieldworker might repeat or play back a constituent while looking at the transcript with the consultant. Consultants might then do one of four things:

- Repeat what the fieldworker has said, but with better pronunciation and with a translation
- Repeat a subset of what the fieldworker has said, and provide a translation of that selection
- Repeat what the fieldworker has said, but add more to the transcribed material to complete a constituent, and then provide a translation of that
- Be confused and unable to proceed with the task because the repeated portion is badly mispronounced or is not a constituent

Native speaker intuitions on constituency, as reflected in one of these four responses, are invaluable at this stage of analysis.

Be aware that while speakers may have a good feel for constituent boundaries, their intuitions will not be consistently reliable. Pauses, intonation, and verb-final morphology can be used by the fieldworker as clues to guess where a constituent or sentence might end, or what the boundaries of a subordinate clause or final clause ending are. Often, prosodic cues will line up with constituent endings except when the speaker skews this alignment for special effect. For example, the speaker may pause before the last phrase of a clause to evoke suspense. See Woodbury (1985) for a discussion of default and skewed alignment between prosody and syntactic constituents. In my [de Reuse's] experience, a Hän (Athabascan of Alaska) speaker's feel for where a sentence ended often (over 80% of the time) coincided with my guesses. But working on the same text with several Hän speakers often resulted in speaker disagreements on sentence boundaries. See Himmelmann (2006:258–270) for challenges in transcribing discourse-level units and determining the boundaries between them.

Matthewson (2004:383) suggests that speakers should only be asked for translations of sentences, and not of dependent clauses or phrases. This is because translations that purportedly have to do with one clause may actually be relevant for the whole sentence.

The meaning of a clause may only make sense when the whole sentence is taken into consideration, so, unbeknownst to the fieldworker, the translation could include portions of the rest of sentence. The suggestion to translate only whole sentences certainly makes sense for non clause-chaining languages but not for clause-chaining languages. In clause-chaining languages, a series of subordinate clauses is strung together and the final clause in the clause-chain construction occurs with a verb with clause-final finite morphology. There are two problems with asking speakers to translate a whole sentence in clause-chain constructions. First, "the sentence" is not a construct that speakers necessarily recognize. In our experience, some speakers recognize that particular morphological sequences occur clause finally, but do not distinguish between clause-final finite morphology and clause-final subordinating morphology. Second, some speakers see each subordinate clause as a semantic whole and translate that clause as a full sentence, e.g., *After that, he was walking in the forest. After that, he came to a big tree.*' We have also found that whether the clause in question is main or subordinate clause, a speaker may report that "something more must follow".

The fieldworker should be careful about speakers changing the wording of the text. Some speakers want to clean up or "regularize" the text to fit a prescriptive standard. Of course, the fieldworker should note what the consultants say in these instances; however, the transcribed portion should not be deleted until the fieldworker has the opportunity to check whether the recording matches the original transcription or the "corrected" version. Also, free translations are useful in completing morpheme analysis but at times can be too literary to be of help (Kibrik 1977).

Again, we note that it is useful to have free translations from at least two speakers, and more if possible. The free translations differ slightly from person to person because each individual approaches the text based on their individual history with that text, and each speaker has a different talent and appreciation for cultural details (Scollon 1979:13). I [Chelliah] have found that urban and rural Manipuri speakers have different understandings of details in traditional stories. Some urban speakers I have worked with have forgotten details about weaving and types of cloth; a rural speaker was able to supply these details and clarify lexical items in a story. In terms of speaker talent, we recommend that, if possible, one of the speakers who helps with the free translation be fluent in English (assuming that this is the language to be used for academic publications).

In the case of older narratives that speakers may have memorized or that may exist in manuscript form, speakers may be able to provide free translations but may hesitate to give word-for-word or clause translations. For example, there are no linguistically annotated pre-twentieth century Manipuri texts, but there are several texts from that period for which free translations exist (Chelliah and Ray 2002). When I [Chelliah] tried to find speakers to help with translations of these texts, they reported the following problems:

- The words are archaic and the consultant does not know what they mean.
- The use of words is figurative, so it is not the consultant's place to pin down the literal and figurative meanings, and someone with more authority would have to do that.
- The texts are sacred and should be worked on by community-sanctioned scholars only.

As a result, it is impossible for a lone linguist to obtain full translations of these texts; a long-term community effort is required. See also Evans and Sasse (2007:227) on difficulties in translating esoteric and archaic material.

13.4.4.7 Morphological Analysis

Details on morphological typology and methods of morphological analysis and elicitation are given in Chapters 11 and 12. Morphological analysis is one of the most difficult and rewarding aspects of annotating textual data. The context afforded by texts helps in determining the meaning of morphemes. However, even years after working on a language, the fieldworker may still need to revise or refine early analyses. It is important, then, to use data management and annotation software that will allow the researcher to keep track of the way a morpheme is glossed and should permit changing that gloss globally if necessary. At the time of the writing of this book, popular software for text annotation is SIL's TOOLBOX or SIL's Fieldworks Language Explorer (FLEx). It is not recommended that interlinear glosses be entered in a word processor because even after glosses are aligned and translations are typed in, the data in the resulting file cannot be automatically searched (e.g. using a concordance program) or linked to other components in the documentation project (e.g. word lists). Furthermore, for archiving purposes the fieldworker should use software whose output is maximally portable and does not require proprietary software to read or process. Thus, it makes more sense to invest a few days learning how to use a program like FLEx, and utilizing its many features such as linking transcription to audio, automatic gloss fill-in, lexicon generation, and database searches.[7] It is true that digitized recordings and text annotation software do not, in themselves, allow the linguist to improve the quality of text annotation (Evans and Dench 2006:25). However, because annotation programs allow the linguist to improve on annotation gradually as their understanding of the target language increases, the annotations become richer and more accurate.

As discussed by Anna Margetts (2009), even the most basic syntactic analysis – for example, checking on constituent boundaries – hearing the utterance along with studying a transcription of the utterance can be helpful. Therefore, software which allows for quick retrieval of the sound file associated with a transcript is necessary. Alignment of source audio and video material to transcription is made possible by software such as ELAN or TRANSCRIBER. See Andrew Margetts (2009).

Standardized terminology and abbreviations should be used for morpheme analysis as far as possible. Some models are the Leipzig Glossing Rules (Comrie et al. 2008), the EUROTYP conventions (König et al. 1993), or E-MELD GOLD ontology (E-Meld 2006a). A fieldworker will find it helpful to maintain a list of abbreviations and glossing conventions handy in hard copy even though this vocabulary will be subject to some modification.

[7] See http://www.sil.org/computing/catalog/show_software.asp?id=79

13.4.4.8 The Final Product

The final product of a text collection is a set of annotated texts. Woodbury (2005) argues for "thick translation" of texts as part of the documentary record of language so that annotations include several levels of information including free translations; sentence-by-sentence translations; glossing by native speaking linguists; and literary translations that are repeatedly improved. A fleshed-out interlinear annotation may have the following structure:

- Practical orthography
- Close phonetic transcription
- Phonemic transcription
- Morpheme-by-morpheme boundaries
- Morpheme gloss
- Word gloss
- Constituent gloss
- Free translation of clause
- Ethnographic notes
- Gesture notes
- Comments

A comprehensive discussion of the different tiers is given in Schultze-Berndt (2006). The final format of a collection of texts will be determined by its purpose. If meant as a community resource, the collection might include only transcription in a practical orthography, word-for-word translation, and free translation. On the other hand, if the text is for syntactic or morphological study, all the glosses might be included, but the narrow phonetic transcription and/or the orthographic representation omitted. If the translation is of a signed language, then the tiers may be as given in Zaefferer (2006:125), which would include phonological representation of mouthings, representation of weak and strong hand signs, and morphological representations.

Cultural and ethnographic notes are necessary for the proper translation and comprehension of texts. The use of a particular grammatical construction or conversational response may be explained culturally. See, for example, Hill's (2006:616) discussion of Sapir's explanation of the Takelma inferential. Ethnographic details help to "gear language documentation towards a holistic perspective" (Widlock 2004:5). This can involve [identifying / sorting out / sifting through / distinguishing / unraveling / disentangling] different layers of information collected [within / woven throughout] one fieldwork session, or it can involve making connections between data collected over the course of several sessions, or between new data and earlier texts or archived material. So one traditional story might include the names of traditional baskets, and this information might be available in list form in another part of the archive. Likewise, a text might include basket names, but also words for species of fish or place names. Metadata fields for ethnographic information should be included for texts.

In addition to speaker information, standard information about the transcriber and translator should be added to the metadata. These speakers may have command of a different dialect and will certainly provide examples and comments that should also be included in the corpus.

Like a grammar or dictionary, a collection of texts is never done. There is always room to improve an analysis or to add more information in the form of another annotation tier (Simpson 2007). We should also not assume that the annotation process will become better or faster as new recording devices or text-analysis software become available (Evans and Dench 2006:25): the basic challenges of transcription and translation remain. In addition, annotations evolve over a period of time, with contributions from a number of people (Schultze-Berndt 2006:217).

Finally, a collection of annotated texts may not be particularly valuable to community members, since most of the information is intended for use by linguists. Nathan (2006:368–369) characterizes software such as TOOLBOX as representative of a 'thin interface'. That is, while TOOLBOX makes data structures transparent, and easy to search and extract information from, it does not prepare data for [dissemination through/sharing with] the native-speaker community. As Nathan characterizes it, for data mobilization, 'thick interfaces' are required; that is, materials are should be designed to take user needs into consideration, and should make effective use of hypertext and multimedia. The final product should reflect community input and collaboration; multimedia design should take usability by community members into consideration (Nathan 2006:370).

13.4.5 Recording and Analyzing Conversations

Conversations are difficult to record. From a practical point of view, the fieldworker must be concerned about microphone placement so all speakers are recorded equally well. From an ethical standpoint, everyone involved must be willing to be recorded and must agree that the recording can be used in analysis and publication. From a transcription and analysis perspective, it is difficult – even with the help of native speakers – to transcribe fast, natural speech with overlaps, interruptions, and sudden changes of topic. Finally, the observer's paradox is in full play, as it is difficult for a conversation to be fully "natural" when the recorder in on. However, linguists have found reasonable solutions to each of these problems, so there is no longer any excuse for not including conversational data in analysis. Again, the sad exception is for languages where there are not enough speakers remaining for natural conversations to take place, or when speakers do not recall enough of the language to carry on a conversation. See, for example, the account of the abilities of the last and only speaker of Tunica, a language of Louisiana, by Haas (1941:9–10).

For microphone placement and other technical hints on recording conversations, see Jackson (1987) and E-MELD (2005). A range of recording situations can arise, from optimal to emergency. In the best case scenario, every speaker has a designated microphone which may be stand-alone, a headset, or a lapel microphone, and the recorder is inconspicuously placed. In an emergency situation – for instance, a rare opportunity to record a moribund language – there might be one microphone and several speakers. The fieldworker must improvise to get the best recording of the ongoing conversation. One possibility is to create a tight circle of speakers, sit at the center of a circle and hold the microphone towards the current speaker. In my [Chelliah's] experience, if the fieldworker does not make eye contact with speakers, and especially if he or she appears not to be following the conversation closely, the speakers will ultimately ignore him or her. If a strong unidirectional microphone is being used, it should pick up a clear signal. This is an absolutely last resort option but it is better to try to capture an interaction than to not try at all. See Cukor-Avila (2006) for other suggestions on recording conversations.

A range of agreements can be reached with speakers about how the recorded conversations can be used.[8] Speakers may request that recorded conversations be erased. They may be willing to help transcribe and translate the conversation themselves, but may request that the conversation not be played back to anyone else. In this case, the data might be used for analysis and publication, if traces of individual identities are removed. Speakers will be a bit puzzled about why conversations are of interest to the researcher, so the fieldworker must think of an appropriately phrased explanation. Permissions should be discussed with speakers after the recording as well as before, since it is only after the recording that speakers know what the topic of the conversation is. If possible, the fieldworker should record these permissions as well.

As for transcription and analysis of narrative, there are several easy-to-use programs, such as TRANSCRIBER, that allow the researcher to:

- Transcribe and label the speech of interlocutors
- Represent the interruptions and overlaps of turns by providing a different transcription tier for each speaker linked to a sound file
- Allow for the speech signal to be slowed down and replayed to help with transcription
- Facilitate the time-alignment of video, audio, transcription, and analysis, with the additional use of programs like ELAN

Useful transcript symbols for conversation transcription and analysis are in Edwards and Lampert (1993) and in the transcription module of Emanuel A. Schegloff's homepage.[9]

13.4.6 Working with an Existing Corpus of Texts

A previously analyzed corpus should be incorporated into a fieldwork project, keeping in mind the following issues: confusion of form and function in morpheme analysis and glossing; consistency in transcription, including indication of clause

[8] For legal issues see Mark Liberman's discussion online at http://www.ldc.upenn.edu/exploration/expl2000/papers/liberman/liberman.html

[9] http://www.sscnet.ucla.edu/soc/faculty/schegloff/

and phrase boundaries; and reliability of free translations. For detailed discussion of philological issues, see Section 5.2.

The morphological analyses of previous fieldworkers should be taken as a suggestion, and not accepted at face value, since it is often impossible to know the intention, level of linguistic sophistication, or effort put into text analysis by someone else. Furthermore, unless discussed in detail by the annotator, one can never be quite sure what his or her grammatical category labels mean. Mosel (2006:51) notes a common terminological blurring with use of the term *adverb*: it can be used to refer to the word class *adverb*, or it can be used to refer to the syntactic function *adverb*, lexically expressed using an adjective or prepositional phrase. Another example is when an eyewitness evidential is glossed as a marker of past tense, because it is found primarily in past tense contexts. (This is motivated by the fact that an event that has been witnessed has already occurred.) One type of evidence that form and function have been confused by a previous fieldworker is when two morphemes have been assigned the same gloss; in the example just mentioned, for instance, both an evidential marker and a tense marker – that is, two different morphemes – might be glossed ‘past tense’. This confusion is analytically useful because it reveals that the meaning of past (completion and remoteness) are indicated by the morphemes in question. Careful questioning and investigation of the textual context can show how the morphemes are different.[10]

It is often difficult to decipher the transcription conventions used by an earlier linguist. Ivy Doak (p.c.) notes, in her work on Coeur d’Alene (Interior Salish of Idaho) using the texts of Reichard that:

> the initial problem with working with Reichard’s [Coeur d’Alene] texts has been understanding her transcription system and the phonology of the language. Things like vowel harmony, glottal releases transcribed two different ways, and schwas transcribed with three different symbols made it difficult to muddle through the texts. A second problem, once past the phonology, is in figuring out the breaks that Reichard indicates with periods and paragraph markers. Her periods are sparse, and often do not correspond to syntactic/intonational units that would be evident in speech from my consultants. Some period phrases include three or more predicates that do not seem to be part of anything like a serial verb type of construction; other period phrases seem syntactically incomplete (some periods she has removed to join an argument with a predicate).

Finally, earlier fieldworkers’ free translations might be too idiomatically aligned with the contact language – so much so that the structure of the target language can be obscured. Again, Ivy Doak (p.c.) comments on Reichard’s unpublished Coeur d’Alene texts, from 1942, for which she finds the translations to be “very English” in comparison to modern stories Ivy Doak has recorded. She says that Reichard’s:

> free translations are quite readable, but do not reflect the structure of the sentences even as she has them analyzed. This is great for getting the plot clear, but not so good for looking at stylistics, story-telling skill, and story structure.

[10] See Matthewson (2004) for a slightly different point of view.

On a similar point, Evans and Sasse (2007:265) say about parallel translations that[11]:

> as useful as they are…parallel texts only address standardized, universal stories, and fail to explore what is culture-specific, either in terms of stories or in terms of lexical items. Parallel Bible or other corpora may tell us how to say 'arise!' or 'Cain fought with Abel'. But we will not encounter the whole subworld of lexical particularities that make a language unique, such as the Dalabon [Australian] *dalabborrord* 'place on a tree where the branches rub together, taken advantage of in sorcery by placing something that has been in contact with the victim, such as clothes, in such a way that it will be rubbed as the tree blows in the wind, gradually sickening and weakening the victim'. The thousands of fascinating words of this type are simply bracketed out from traditions of parallel translation.

References

Albert, Ethel. 1972. 'Rhetoric', 'logic', and 'poetics' in Burundi: culture patterning of speech behavior. In *Directions in Sociolinguistics. The Ethnography of Communication*, ed. by John J. Gumperz and Dell Hymes, 72–105. New York: Holt, Rinehart & Winston.

Ameka, Felix K. 2006. Real descriptions: Reflections on native speaker and non-native speaker descriptions of a language. In *Catching Language: The Standing Challenge of Grammar Writing*. (Trends in Linguistics. Studies and Monographs 167.) ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 69–112. Berlin: Mouton de Gruyter.

Amery, Rob. 2009. Phoenix or relic? Documentation of languages with revitalization in mind. *Language Documentation and Conservation* 3(2):138–148. Online: http://nflrc.hawaii.edu/ldc/http://hdl.handle.net/10125/4436.

Amha, Azeb, and Gerrit Dimmendaal. 2006. Converbs in an African perspective. In *Catching Language: The Standing Challenge of Grammar Writing*. (Trends in Linguistics. Studies and Monographs 167.) ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 393–440. Berlin: Mouton de Gruyter.

Austin, J. L. 1962. *How to Do Things with Words*. Cambridge, MA: Harvard University Press.

Austin, Peter K. 2006. Data and language documentation. In *Catching Language: The Standing Challenge of Grammar Writing*. (Trends in Linguistics. Studies and Monographs 167.) ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 87–112. Berlin: Mouton de Gruyter.

Bach, K. and R. M. Harnish. 1979. *Linguistic Communication and Speech Acts*. Cambridge, MA: MIT Press.

Boas, Franz. 1917. Introduction. *International Journal of American Linguistics* 1:1–8. [Reprinted in F. Boas. 1940. *Race, Language, and Culture*. 199–210. New York: The Macmillan Company.]

Bow, C., B. Hughes, and Steven Bird. 2003. Towards a general model of interlinear text. In *Proceedings of the E-MELD Language Digitization Project: Workshop on Digitizing and Annotating Texts and Field Recordings*. Online: http://emeld.org/workshop/2003/bowbadenbird-paper.html

Burt, Susan Meredith. 2008. Contact Pragmatics: Requests in Wisconsin Hmong. *Proceedings of the 2007 Meeting of the Southeast Asian Linguistics Society*, 62–75. Online: http://www.jseals.org/.

Caldecott, Marion G., and Karsten A. Koch. n.d. We have ways of making you talk: Eliciting conversations in FN (First Nations) languages. Online: http://www.uleth.ca/conferences/wscla12/Abstracts/Caldecott&Koch.pdf.

[11] Translations that place the target language and the free translation in a format so that they can be easily compared, say in side-by-side columns, are called parallel translations.

Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In *Subject and Topic*, ed. by Charles N. Li, 27–55. New York: Academic.

Chafe, Wallace L., ed. 1980. *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. Norwood, NJ: Ablex.

Chelliah, Shobhana L. 2001. The role of text collection and elicitation in linguistic fieldwork. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 152–165. Cambridge: Cambridge University Press.

Chelliah, Shobhana L., and Sohini Ray. 2002. Early Meithei Manuscripts. In *Medieval Tibeto-Burman Languages:* Proceedings of a Symposium held in Leiden, June 26, 2000, at the 9th Seminar of the International Association of Tibetan Studies, ed. by Christopher I. Beckwith, 59–71. Leiden: E.J. Brill.

Cholin, Jochen. 2004. Video recording in the field. *Language Archives Newsletter* 1(4):5–8. Online: http://www.mpi.nl/LAN.

Comrie, Bernard, Martin Haspelmath, and Balthasar Bickel. 2008. Conventions for interlinear morpheme-by-morpheme glosses. Online: http://www.eva.mpg.de/lingua/resources/glossing-rules.php.

Coupe, Alec R. 2007. *A Grammar of Mongsen Ao* (Mouton Grammar Library 39). Berlin and New York: Mouton de Gruyter.

Crowley, Terry. 2007. *Field Linguistics. A Beginner's Guide*. Oxford: Oxford University Press.

Cruse, D. Alan. 1986. *Lexical Semantics*. Cambridge: Cambridge University Press.

Cruse, D. Alan. 2004. *Meaning in Language: An introduction to Semantics and Pragmatics*. 2nd ed. Oxford: Oxford University Press.

Cukor-Avila, Patricia. 2006. Researching Naturally Occurring Speech. *Encyclopedia of Languages and Linguistics*. 2nd Edition, 556–563. Oxford: Elsevier.

Dixon, R. M. W. 2007. Field Linguistics: a minor manual. *Sprachtypologie und Universalienforschung* (Focus on Linguistic Fieldwork, ed. by Alexandra Y. Aikhenvald) 60(1):12–31.

Dixon, R. M. W. 2010. *Basic Linguistic Theory. Volume 1 Methodology.* Oxford: Oxford University Press.

Dorian, Nancy C. 2001. Surprises in Sutherland: linguistic variability amidst social uniformity. In *Linguistic Fieldwork*, ed. by Paul Newman and Martha Ratliff, 133–151. Cambridge: Cambridge University Press.

Dwyer, Arienne, and Ulrike Mosel. 2001. *Metadata Description Recommendations: General; Content.* DOBES Technical Reports 6a.1, 6b.2 (15.03.01).

Edwards, Jane A., and Martin D. Lampert, eds. 1993. *Talking Data: Transcription and Coding in Discourse Research*. Hillsdale, NJ: Lawrence Erlbaum Associates.

E-MELD. 2005. *E-MELD School of Best Practice: Hardware Recommendations – Practical Considerations*. Online: http://emeld.org/school/toolroom/hardware/audio-standard3.html.

E-MELD. 2006a. *E-MELD School of Best Practice: Ontologies*.Online: http://emeld.org/school/classroom/ontology/index.html.

E-MELD. 2006b. *E-MELD School of Best Practice: Filming in the Field*. Online: http://emeld.org/school/classroom/video/field.html.

Enfield, Nicholas J. 2006. Heterosemy and the grammar-lexicon trade-off. In *Catching language: The Standing Challenge of Grammar Writing*. (Trends in Linguistics. Studies and Monographs 167.) ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 297–320. Berlin: Mouton de Gruyter.

Evans, Nicholas, and Alan Dench. 2006. Introduction: Catching language. In *Catching language: The Standing Challenge of Grammar Writing*. (Trends in Linguistics. Studies and Monographs 167.) ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 1–39. Berlin: Mouton de Gruyter.

Evans, Nicholas, and Hans-Jürgen Sasse. 2007. Searching for Meaning in the Library of Babel: Field semantics and problems of digital archiving. *Archives and Social Studies: A Journal of Interdisciplinary Research* 1.0:260–320. Online: http://socialstudies.cartagena.es/index.php?option+com_content&task=view&id=38&Itemid=33

Foley, William A. 2002. Field methods. In *The Linguistics Encyclopedia*, 2nd edition, ed. by Kirsten Malmkjær, 131–137. London and New York: Routledge.

Fox, Barbara A. 1987. *Discourse Structure and Anaphora.* Cambridge: Cambridge University Press.

Francis, Norbert, and Pablo Rogelio Navarrete Gómez. 2009. Documentation and language learning: separate agendas or complementary tasks? *Language Documentation and Conservation* 3(2):176–191. Online: http://nflrc.hawaii.edu/ldc/http://hdl.handle.net/10125/4438.

Franchetto, Bruna. 2006. Ethnography in language documentation. In *Essentials of Language Documentation*, ed. by Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel, (Trends in Linguistics. Studies and Monographs 178.), 183–21. Berlin and New York: Mouton de Gruyter.

Geluykens, Ronald. 1994. *The Pragmatics of Discourse Anaphora in English: Evidence from Conversational Repair*. Berlin: Mouton de Gruyter.

Grenoble, Lenore. 2006. Documenting pragmatics and endangered languages. In *Language Documentation and Description* Volume 4, 145–162, ed. by Peter K. Austin. London: SOAS, Hans Rausing Endangered Languages Project.

Grimes, Joseph E. 1995. *Language Survey Reference Guide.* Dallas, TX: Summer Institute of Linguistics.

Haas, Mary R. 1941. *Tunica*. New York: J. J. Augustin.

Healey, Alan, ed. 1975. *Language Learner's Field Guide*. Ukarumpa, Papua New Guinea: Summer Institute of Linguistics.

Hellwig, Birgit. 2006. Field semantics and grammar-writing: Stimuli-based techniques and the study of locative verbs. In *Catching Language: The Standing Challenge of Grammar Writing*. (Trends in Linguistics. Studies and Monographs 167.) ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 321–358. Berlin: Mouton de Gruyter.

Hill, Jane H. 2006. Writing culture in grammar in the Americanist tradition. In *Catching Language: The Standing Challenge of Grammar Writing*. (Trends in Linguistics. Studies and Monographs 167.) ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 609–628. Berlin: Mouton de Gruyter.

Himmelmann, Nikolaus P. 2006. The challenges of segmenting spoken language. In *Essentials of Language Documentation*. (Trends in Linguistics. Studies and Monographs 178.) ed. by Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel, 253–274. Berlin and New York: Mouton de Gruyter.

Hoijer, Harry. 1958. Native Reaction as a Criterion in Linguistic Analysis. In *Proceedings of the Eighth International Congress of Linguists*, ed. by Eva Sivertsen, 573–583. Oslo: Oslo University Press.

Jackson, Bruce. 1987. *Fieldwork*. Urbana, IL and Chicago, IL: University of Illinois Press.

Karttunen, Lauri. n. d. Presuppositional Phenomena. Unpublished manuscript. University of Texas, Austin, TX.

Kibrik, Aleksandr E. 1977. *The Methodology of Field Investigations in Linguistics. (Setting up the Problem.)* (Janua Linguarum, Series Minor, 142.) The Hague: Mouton.

König, Ekkehard, (with Dik Bakker, Östen Dahl, Martin Haspelmath, Maria Koptjevskaja-Tamm, Christian Lehmann, and Anna Siewierska). 1993. *EUROTYP Guidelines*. Berlin and Strasbourg: European Science Foundation Programme in Language Typology. Online: http://www-uilots.let.uu.nl/ltrc/eurotyp/index.htm.

Kroeger, Paul. 2004. *Analyzing Syntax. A Lexical-Functional Approach*. Cambridge: Cambridge University Press.

Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: Pennsylvania University Press.

Lehmann, Christian. 1982. Directions for interlinear morphemic translations. *Folia Linguistica* 16:199–224.

Levinson, Steven. 1983. *Pragmatics*. Cambridge: Cambridge University Press.

Lyons, John. 1977. *Semantics.* Vol. 1. Cambridge: Cambridge University Press.

McConvell, Patrick. 2003. Multilingual Multiperson Multimedia: Linking Audio-Visual with Text Material in Language Documentation. *Researchers, Communities, Institutions, Sound Recordings*, ed. by Linda Barwick, Allan Marett, Jane Simpson and Amanda Harris. Sydney: University of Sydney. Online: http://ses.library.usyd.edu.au/bitstream/2123/1429/1/McConvellrev1.pdf

McConvell, Patrick. 2007. Video – A linguist's view. (A reply to David Nathan.) *Language Archives Newsletter* 10:2–3. Online: http://www.mpi.nl/LAN.

Margetts, Andrew. 2009. Using toolbox with media files. *Language Documentation and Conservation* 3(1):51–86. Online: http://nflrc.hawaii.edu/ldc/http://hdl.handle.net/10125/4426.

Margetts, Anna. 2009. Data processing and its impact on linguistic analysis. *Language Documentation and Conservation* 3(1):87–99. Online: http://nflrc.hawaii.edu/ldc/http://hdl.handle.net/10125/4425.

Matthewson, Lisa. 2004. On the methodology of semantic fieldwork. *International Journal of American Linguistics* 70(4):369–415.

Mayer, Mercer. 1969. *Frog Where Are You?* New York: Penguin Putnam.

Milroy, Lesley. 1987. *Observing and Analysing Natural Language. A Critical Account of Sociolinguistic Method.* (Language in Society 12). Oxford: Basil Blackwell.

Mirzayan, Armik. 2008. A Preliminary study of same-turn self-repair initiation in Wichita conversation. In *Lessons from Documented Endangered Languages*. (Typological Studies in Language 78.) ed. by David K. Harrison, David S. Rood, and Arienne Dwyer, 317–354. Amsterdam and Philadelphia: John Benjamins.

Mosel, Ulrike. 2004. Inventing communicative events: Conflicts arising from the aims of language documentation. *Language Archives Newsletter* 1(3):3–4. Online: http://www.mpi.nl/LAN.

Mosel, Ulrike. 2006. Sketch Grammar. In *Essentials of Language Documentation*. (Trends in Linguistics. Studies and Monographs 178.) ed. by Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel, 301–309. Berlin and New York: Mouton de Gruyter.

Nathan, David. 2006. Thick interfaces: Mobilizing language documentation with multimedia. In *Essentials of Language Documentation*. (Trends in Linguistics. Studies and Monographs 178.) ed. by Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel, 363–379. Berlin and New York: Mouton de Gruyter.

Nathan, David. 2007. Digital video in documentation and archiving. *Language Archives Newsletter* 9:3–4. Online: http://www.mpi.nl/LAN.

Payne, Thomas E. 1997. *Describing Morphosyntax: A Guide for Field Linguists*. Cambridge: Cambridge University Press.

Rice, Keren. 2006. Let the language tell its story? The role of linguistic theory in writing grammars. In *Catching Language: The Standing Challenge of Grammar Writing*. (Trends in Linguistics. Studies and Monographs 167.) ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 235–268. Berlin: Mouton de Gruyter.

Saeed, John I. 2009. *Semantics*, 3rd edn. Malden, MA: Wiley-Blackwell.

Sapir, Edward. 1921. *Language: An Introduction to the Study of Speech*. New York: Harcourt Brace.

Schiffrin, Deborah. 1994. *Approaches to Discourse*. (Blackwell Textbooks in Linguistics.) Oxford and Cambridge: Blackwell.

Schultze-Berndt, Eva. 2006. Linguistic Annotation. In *Essentials of Language Documentation*. (Trends in Linguistics. Studies and Monographs 178.) ed. by Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel, 214–251. Berlin and New York: Mouton de Gruyter.

Scollon, Ronald. 1979. *The Context of the Informant Narrative Performance: From Sociolinguistics to Ethnolinguistics at Fort Chipewyan, Alberta*. (National Museum of Man, Mercury Series, Canadian Ethnology Service Paper No. 52.) Ottawa: National Museums of Canada.

Silko, Leslie Marmon. 2001. Language and Literature from a Pueblo Indian Perspective. In *Nothing but the Truth. An Anthology of Native American Literature*, ed. by John L. Purdy and James Ruppert, 159–165. Upper Saddle River, NJ: Prentice-Hall.

Simpson, Jane. 2007. When is a linguist's work done and dusted? Online: http://blogs.usyd.edu.au/elac/2007/04/when_is_a_linguists_work_done.html.

Sutton, Peter, and Michael Walsh. 1979. *Revised Linguistic Fieldwork Manual for Australia*. Canberra: Australian Institute of Aboriginal Studies.

Tedlock, Dennis. 1983. *The Spoken Word and the Work of Interpretation*. Philadelphia: University of Pennsylvania Press.

Widlok, Thomas. 2004. Ethnography in language documentation. *Language Archives Newsletter* 1(3):4–6. Online: http://www.mpi.nl/LAN/.

Wittenburg, Peter. 2007. Video – A technologist's view. (A reply to David Nathan.) *Language Archives Newsletter* 10:3–5. Online: http://www.mpi.nl/LAN.

Woodbury, Anthony C. 1985. The functions of rhetorical structure: a study of Central Alaskan Yupik Eskimo discourse. *Language in Society* 14:153–190.

Woodbury, Anthony C. 2005. On thick translation in linguistic documentation. Paper presented at the *ELAP Workshop: Meaning and Translation in Language Documentation*. London: Hans Rausing Endangered languages Project, School of Oriental and African Studies.

Yamamoto, Akira Y. 1998. Linguists and Endangered Language Communities. Issues and Approaches. In *Studies in Endangered Languages: Papers from the International Symposium on Endangered Languages*, Tokyo, November 18–20, 1995. (International Clearinghouse for Endangered Languages linguistic studies 1.) ed. by Kazuto Matsumura, 213–252. Tokyo: Hituzi Syobo.

Zaefferer, Dietmar. 2006. Realizing Humboldt's dream: Cross-linguistic grammatography as database creation. In *Catching language: The Standing Challenge of Grammar Writing*. (Trends in Linguistics. Studies and Monographs 167.), ed. by Felix K. Ameka, Alan Dench, and Nicholas Evans, 113–135. Berlin: Mouton de Gruyter.

Index

H

J