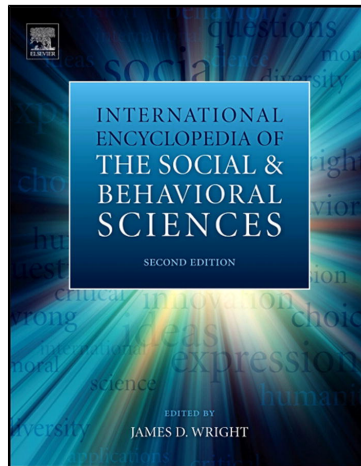


**Provided for non-commercial research and educational use only.
Not for reproduction, distribution or commercial use.**

This article was originally published in the *International Encyclopedia of the Social & Behavioral Sciences*, 2nd edition, published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues who you know, and providing a copy to your institution's administrator.



All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

<http://www.elsevier.com/locate/permissionusematerial>

From Berez, A.L., 2015. Linguistic Fieldwork. In: James D. Wright (editor-in-chief), *International Encyclopedia of the Social & Behavioral Sciences*, 2nd edition, Vol 14. Oxford: Elsevier. pp. 183–189.

ISBN: 9780080970868

Copyright © 2015 Elsevier Ltd. unless otherwise stated. All rights reserved.

Elsevier

Linguistic Fieldwork

Andrea L. Berez, University of Hawai'i at Mānoa, Honolulu, HI, USA

© 2015 Elsevier Ltd. All rights reserved.

This article is a revision of the previous edition article by U. Mosel, volume 13, pp. 8906–8910, © 2001, Elsevier Ltd.

Abstract

This article describes the undertaking of linguistic fieldwork, in which linguists study languages as they are spoken in a naturally occurring domain, rather than in a laboratory or via self-introspection. This article presents the history of linguistic fieldwork, the rise of language documentation as a major paradigm for fieldwork, fieldwork research methods, ethics and community relations, and data management issues.

What Is Linguistic Fieldwork?

The traditional understanding of *linguistic fieldwork* involves a trained linguistic scientist traveling to a remote location to live in a speech community in order to study, describe, and theorize about that community's language or the human language capacity in general. In describing the nature of prototypical fieldwork, Hyman (2001) lists three features that linguistic fieldwork exhibits. The first of these is *distance*, meaning that the research usually takes place at some distance from one's home ("That is, 'fieldwork on Kikuyu' on one's vitae becomes synonymous with 'I did fieldwork on Kikuyu in Kenya'" (Hyman, 2001, p. 16)). The second of these is *exoticism*, meaning that the language under study is usually a minority language, an understudied language, an endangered language, etc. Hyman's third characteristic of prototypical fieldwork is *duration*, meaning that the research must take place over an extended period of time, with a commitment to spending many months in the field and years investigating the language.

While these traditional characteristics of fieldwork are considered by some to be the hallmarks of true linguistic fieldwork (e.g., Crowley, 2007; Samarin, 1967), in the last decade other less-prototypical scenarios are becoming increasingly accepted and even encouraged as valid examples of linguistic fieldwork. For example, fieldwork may take place in a speech community that is living in a diaspora situation away from its homeland, perhaps in a large urban setting like New York City or London. Fieldwork may also take place in a culturally defined subcommunity of the majority language community; for example, Eckert's (1989, 2000) and Bucholtz's (2011) work on language use among American high school students. Fieldwork may even be considered to start in the college classroom in a linguistics field methods course or by working with a lone individual, although these scenarios usually lead to the more traditional model of time spent in-community by the researcher later, in order to observe the language in use in a naturalistic setting.

What all these definitions have in common is that in fieldwork, the linguist attempts to study a language by observing it in one or more of its naturally occurring domains of use. While the notion of a *domain of use* can vary from a remote village where the language has been spoken for generations, to an urban marketplace, to a religious ceremony, to a staged conversation between two fluent speakers who are

far from the home community, fieldwork is distinct from other linguistic research methods in that it strives to collect data by working with people who speak the language under study (as opposed to, say, introspective or corpus-based methods).

The products or outcomes expected from linguistic fieldwork are also changing. Previously the expected products of fieldwork were limited to those that primarily served linguistic science, whether descriptive (i.e., the so-called 'Boasian trilogy' of a dictionary, a descriptive grammar, and a collection of texts) or theoretical (examples from an 'exotic' language to prove or disprove theories) in aim. Today, however, the products of fieldwork can reach a much broader audience. The focus since Hale et al. (1992) on the endangered language crisis has led to an increase in the production of language teaching materials and revitalization programs as a goal of fieldwork, as well as increased attention to capacity building activities in language communities. The emerging discipline of language documentation, described below (and see, e.g., Gippert et al., 2006; Himmelmann, 1998, 2012; Grenoble and Furbee, 2010; Woodbury, 2011), has also led to the production of annotated digital media corpora consisting of audio and video recordings of language in use, along with time-aligned transcription, translation, and analysis as a central goal of linguistic fieldwork. These documentary corpora are intended to serve as wide an audience as possible, including linguists, language communities, teachers, and others.

Furthermore, the notion of who is a fieldworker is changing. Traditionally, linguistic fieldworkers have been professional linguists, graduate students, or Christian missionaries studying a language in order to facilitate Bible translation. Now, however, it is common to find members of the speech community themselves doing fieldwork; these people may or may not be trained in linguistic science, and may be teachers, clergy, or interested laypersons with a personal stake in the study of their native or heritage language.

In fact, the term *fieldwork* has recently come under scrutiny. This is because for many people who undertake the study of language *in situ*, the traditional notion of 'going to the field' does not hold – these people are often working in their home communities, with relatives or friends serving as language consultants. However, this article continues to use the term *fieldwork* as it is used in the sciences, to distinguish the study of language in its natural environment from that in

a laboratory setting; the terms *fieldworker* and *field linguist* here are used to indicate the person who undertakes such activities, whether or not that person is affiliated with an academic institution.

History of Linguistic Fieldwork

The history of linguistic fieldwork, broadly defined, goes back at least five centuries, when Spanish and Portuguese colonists in Africa, Asia, and the Americas embarked on the study of indigenous languages in the service of exploration and Christian mission work. Philologically grounded interest in the collection of wordlists was also popular in eighteenth- and nineteenth-century Europe among scholars and laypersons alike; see Chelliah and De Reuse (2010) for a full account of the early history of linguistic fieldwork.

Contemporary fieldwork is generally considered to have started with Franz Boas in the late nineteenth and early twentieth centuries (see Rosenblum and Berez, 2010; Sakel and Everett, 2012; Woodbury, 2011). In his anthropological work in indigenous communities along the Pacific Northwest coast of North America, Boas stressed the importance of language as an inextricable component of ethnological studies as a whole. He advocated the collection of texts, especially those written by native speakers, as part of the complete description of a language. This was a sharp turn from previous linguistic fieldwork, which focused on the production of generalizing descriptive grammars and dictionaries or wordlists as the essential products of fieldwork. Boas also became the progenitor of a tradition of American fieldwork-based linguistic inquiry via his arguably most-notable student Edward Sapir, who trained, among others, Mary Haas and Morris Swadesh. Haas then went on to train at least 50 graduate students in linguistic fieldwork through her position at the University of California, Berkeley, and the Survey of California and Other Indian Languages (Turner, 1997).

Turning to other regions, the history of fieldwork on the Aboriginal languages of Australia, summarized in McGregor (2008), can roughly be divided into three periods: an early period from the 1770s until about the 1930s that was marked by the collection of wordlists and a few descriptive grammars by amateurs; a middle period from approximately the mid-1930s until the 1960s, in which professional linguists conducted fieldwork, equipped with new ideas and institutional backing; and a late period from the 1960s to the present, in which the presence of missionary linguists and Aboriginal involvement in fieldwork has expanded rapidly (see Blake, 1981; Capell, 1970; Dixon, 1980; Dixon and Blake, 1979; McGregor, 2004; O'Grady et al., 1966; Wurm, 1972 for more details).

Furthermore, fieldwork in the Pacific region – and especially on languages of the Austronesian family – was facilitated during the second half of the twentieth century in large part by linguists working at the University of Hawai'i at Mānoa. There, the Pacific Lexicography Center and, later, the Pacific and Asian Linguistics Institute, supported fieldwork in languages of Micronesia, Polynesia, and the Philippines (Grace, unpublished).

During the rise of generative linguistics in America in the second half of the twentieth century, fieldwork, while still a popular research methodology in Australia and the Pacific, fell out of favor in the United States (with the exception of research on languages of the Americas as part of the Boasian tradition described above). As Chelliah and De Reuse (2010) note:

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.

Chelliah and De Reuse, 2010, pp. 60–61

The Rise of Language Documentation as a Paradigm for Fieldwork

However, since the early 1990s, attention has once again turned to the value of linguistic fieldwork, spurred in particular by the crisis of language endangerment and extinction (e.g., Austin and Sallabank, 2011; Evans, 2009; Grenoble and Whaley, 1998; Hale et al., 1992; Harrison, 2007; Nettle and Romaine, 2000). Krauss's (in Hale et al., 1992) admonition that linguists not let the discipline "go down in history as the only science that presided obliviously over the disappearance of 90% of the very field to which it is dedicated" (Krauss in Hale et al., 1992, p. 10) has given rise to a new paradigm of linguistic fieldwork called alternately *language documentation* or *documentary linguistics*. (Some suggest that the term *documentary linguistics* situates this new paradigm as a subfield of linguistics (like *sociolinguistics* or *historical linguistics*), while *language documentation* remains agnostic or relegates it to an applied status. Except in direct quotes, for this article I use *language documentation* because it has become the more widely used of the two terms, although I still maintain that the practice is as worthy of theorization as any other subfield of linguistics (cf Woodbury, 2011).)

Himmelman (1998) distinguishes the *documentation* of languages from the *description* of languages, both of which arise from fieldwork:

Much of the work that is labeled 'descriptive' within linguistics comprises two activities, the collection of primary data and a (low-level) analysis of these data. These are indeed two separate activities as shown by the fact that the methods employed in each activity differ substantially. To date, the field concerned with the first activity – called 'documentary linguistics' here – has received very little attention from linguists. [I propose] that documentary linguistics be conceived of as a fairly independent field of linguistic inquiry and practice that is no longer linked exclusively to the descriptive framework.

Himmelman, 1998, p. 161

The rise of language documentation has had a tremendous affect on linguistic fieldwork in the last two decades. Funding agencies have arisen exclusively to fund fieldwork projects

that are documentary in nature (e.g., The Hans Rausing Endangered Language Documentation Programme, the Dokumentation Bedrohter Sprachen of the VolkswagenStiftung, the National Science Foundation/National Endowment for the Humanities Documenting Endangered Languages Program, the Foundation for Endangered Languages, and the Endangered Language Fund). Fieldwork textbooks now discuss documentation methodology (e.g., Bower, 2008), and academic programs to grant postgraduate degrees in language documentation have been developed (e.g., in the Department of Linguistics at the University of Hawai'i at Mānoa and the Endangered Language Academic Programme at the School of Oriental and African Studies of the University of London). Major research initiatives have explored the infrastructure needed to support language documentation (e.g., The Electronic Metastructure for Endangered Languages Data project (EMELD); inNET). Several edited volumes on documentation and documentary fieldwork (e.g., Gippert et al., 2006; Grenoble and Furbee, 2010; Thieberger, 2012) as well as journals and series (*Language Documentation & Conservation*, *Language Documentation and Description*) have been published.

Summarily, the goals of language documentation fieldwork include the creation of a corpus of audio and video records of observable linguistic behavior, the annotation of those records with enough translation, analysis, and metadata that the records will be useful and accessible to a wide audience, and the responsible preservation and dissemination of those records and annotations (see, e.g., Himmelmann, 1998, 2006; Woodbury, 2011 for discussion). Documentation also goes hand in hand with the field of language revitalization, maintenance, and conservation.

Fieldwork Research Methods and Activities

A great deal of work goes into preparing for fieldwork. These include selecting the language of study, finding an appropriate fieldsite, developing a research agenda, getting visas, finding funding, purchasing equipment, and learning as much as possible about the language, culture, and field site.

Once in the field, fieldworkers typically think of their work as occurring in *sessions*, usually 1–3 h of time spent working with one or more speakers of the language (these speakers are usually called *consultants*, *informants*, *teachers*, *assistants*, etc.). A typical day of fieldwork will consist of one or two sessions, plus preparation time before each session and processing time after each session. A range of activities can take place during a session: *elicitation* of lexical, grammatical, and metalinguistic knowledge; *recording* examples of language in use; and *transcription and translation* of those recordings. In reality the division between these activities is not so clearly defined. For example, during transcription and translation work the linguist and consultant may break away to investigate an interesting grammatical phenomenon or to collect lexical data from a particular semantic field.

Elicitation is the direct questioning of the consultant(s) about a particular aspect of the language. Usually elicitation of lexical items (vocabulary) is the first kind of work to take place between a researcher and a consultant; the researcher will ask the consultant for words of simple nouns like body

parts and use this information to understand the phonology of the language in question (see Bower, 2008 for advice on initial phonetic/phonological fieldwork; Ladefoged, 2003 for advanced instrumental fieldwork on phonetics). Lexical elicitation can also be a part of more advanced fieldwork, in service of producing dictionaries (Frawley et al., 2002; Mosel, 2004; Nichols and Sprouse, 2003). Elicitation can also be used to investigate other parts of grammar, including morphology and syntax. The linguist may ask the consultant 'how do you say X in your language?' or she may provide a scenario and ask the consultant how that scenario is usually expressed in the language ('if you were thirsty, how would you ask for something to drink?'). In some cases the linguist may provide constructions in the language and ask the consultant to explain the difference, or the linguist may ask whether a certain construction is grammatical or not. Some consultants can quickly become adept at providing paradigms. Elicitation can also be aided with the use of short video clips of actions or states as stimulus, like those provided online by the Language and Cognition Group at the Max Planck Institute for Psycholinguistics in Nijmegen.

Elicitation, while a useful method for investigating some aspects of language, requires the consultant to reflect consciously on his or her knowledge of the language. This is cognitively a different kind of linguistic behavior from using language for communication, and thus it is unwise for linguists to base their research solely on elicitation. It is very important to also include the *recording of examples of language in use* in a range of genres and modalities – indeed these kinds of data make up the bulk of a language documentation corpus. Examples of language in use include nearly any kind of running speech like narration, conversation, instructional discourse, oration, etc. These kinds of language use may be more or less spontaneous (see Himmelmann, 1998) and may be either staged for the researcher or naturally occurring.

Staged or controlled speech events for recording may include asking a consultant to tell a traditional story in the language, or asking a consultant to talk about a personal experience, either of her own choosing or in response to a question like 'have you ever been in mortal danger?' or 'tell me about your first day of school.' Visual stimuli may also be used; entire research paradigms have grown out of the use of the 'pear film' (Chafe, 1980) and the 'frog story' (Strömquist and Verhoeven, 2004) in fieldwork. (The 'pear film' is a wordless video about a pear farmer, and a boy, and a bicycle, while the 'frog story' is a wordless picture book about a boy, a dog, and a frog (Meyer, 1969). In these research paradigms, consultants are asked to narrate these in their language.) Conversational speech events may also be somewhat staged by providing consultants with a topic for discussion; Bower (2008, p. 123) suggests games like 'what am I thinking of?' and giving directions. Naturally occurring speech events are also potentially recordable and should be included in a well-rounded corpus. These include public events like ceremonies, performances, lectures, radio programs and the like, as well as conversations among consenting participants (not necessarily in the presence of the linguist).

Recordings of language are of little use if they are not accessible to people who do not speak the language, and therefore transcription and translation are key fieldwork activities.

Transcription is the writing (or typing) of the contents of the recording, usually either in the International Phonetic Alphabet or in a locally accepted orthography. Translation involves translating the transcription into another language (usually the language of wider communication and/or a major language like English), which can happen in units of various sizes. Transcriptions can be translated at the level of phrase or sentence (usually referred to as a *free translation*), at the word level, or at the morpheme level. In all cases, transcription and translation involve analysis, of phonology during transcription and of morphology, syntax, semantics, and discourse during translation.

If the field linguist is also a fluent speaker and writer of the language under study, transcription can be a fairly rapid process. However, if the field linguist is not a speaker of the language, then she must work with a consultant, who can repeat the words on the recording slowly and clearly to the linguist, who will write them down. Translation is also a time-consuming task, and the linguist should take care to avoid the range of pitfalls that come with translating between one language and another (Crowley, 2007; Gil, 2001).

Note that no single method is enough. Consultants who are skilled at metalinguistic activities may not be fluent enough or comfortable enough to provide running speech, and fluent speakers are not necessarily adept at introspection. Importantly, a phrase that sounds ungrammatical out of context may indeed be a licit part of a naturalistic speech event, and thus grammaticality judgments can be suspect. In reality, fieldwork involves all research methods described here happening simultaneously and circularly. As the field researcher learns more about the language, she may need to go back and re-elicite, retranslate, or reanalyze previous work.

Ethics and Linguist–Community Relations

As in medical and biological research, linguists are bound to practice ethical research methods during fieldwork. Ethics has been the subject of a great deal of discussion in the last few decades, with special attention to the notion of community collaboration (discussed below) in language documentation. Because they work with human subjects, most linguists are bound by university-based ethics panels, also referred to as Human Subject Committees or Internal Review Boards. Linguistic fieldwork projects must be approved or exempted by such a committee before the researcher can commence work.

However, given that linguistic fieldwork takes place in parts of the world that may be culturally quite diverse, most researchers agree that it is imperative to consider one's ethical responsibilities beyond what the university ethics panel has approved, since practices for, e.g., consent, may differ from those in the researcher's home country. The Linguistic Society of America has provided some guidelines for ethical field-based research:

- Linguists should do everything in their power to ensure that their research poses no threat to the well-being of research participants.
- Research participants have the right to control whether their actions are recorded in such a way that they can be

connected with their personal identity. They also have the right to control who will have access to the resulting data, with full knowledge of the potential consequences.

- Linguists are responsible for obtaining the informed consent of those who provide them with data (regardless of whether and how that consent is documented), for maintaining any expected confidentiality in storing data and reporting findings, and for ensuring that any individual's participation in their research is completely voluntary at every stage. Anonymous observations of public behavior, which often cannot involve consent, should include no information that could inadvertently identify individuals or, where sensitive, the community.
- Linguists should carefully consider whether compensation of some kind is appropriate, be it remuneration for time and effort, or use of their knowledge and expertise to benefit participants or their communities.
- Where feasible, linguists should facilitate participants' access to their research results (Linguistic Society of America, 2009: pp. 2–3, reprinted in Rice, 2012, p. 410).

These guidelines are designed to protect the individual research participant while still being flexible enough to take cultural differences into account. In traditional social science research, anonymity of a participant is often an inviolable component of project design, but in some language communities, speakers may take pride in having their names associated with their utterances; the second bullet point above allows the speaker to make decisions with regard to his or her anonymity. Bullet point three requires that researchers ensure that participants consent to the research activity, but does not require a signature to that effect, since in many communities paperwork is regarded with suspicion (in such cases, verbal consent would be recorded). Bullet point four allows researchers and consultants to determine locally appropriate compensation, to avoid disrupting the local economy or causing offense (McLaughlin and Seydou Sall, 2001).

In addition to the ethical considerations between linguist and individual research participants, fieldwork also requires ethical consideration between the linguist and the speaker community. Although in many cases it is difficult to determine which individuals or bodies best represent the community, it is now well accepted that fieldworkers must strive to also take the needs, interests, and expertise of the speakers themselves into consideration when designing and conducting field-based projects.

In recent years the discussion about linguist–community relations in fieldwork and language documentation has tended to embrace a collaborative model (Cameron et al., 1992; Czaykowska-Higgins, 2009; Dorian, 2002; Dwyer, 2006; Gerds, 1998; Grinevald, 1998, 2003; Hale, 2001; Hill, 2002; Penfield et al., 2008; Rice, 2006, 2011, 2012; Stebbins, 2003; Yamada, 2007), in which the linguist works cooperatively with the community to mutually meet the language goals of the researchers and the speakers simultaneously. At the very least, the linguist integrates into her research plans to return something of potential linguistic or educational value back to the community while also working toward her own research goals. For example, a field linguist writing a grammar may also use some of the information she collects to create language

teaching materials for teachers to use in local language maintenance programs, or she may install digital 'jukeboxes' of language recordings of traditional narratives on a computer in the local library for all to listen to. Other publications on collaborative research support a much more proactive and equitable relationship between the researcher and the community. Czaykowska-Higgins (2009), for example, extends the notion of *community-based research* (Cameron et al., 1992; Strand et al., 2003) to field linguistics, advocating research that is conducted *for, with, and by* community members. In this model, members of the speech community have an equal say in the design, implementation, and ownership of a fieldwork project:

In its fullest form, Community-Based Language Research involves training members of the language-using community to do the research themselves, and can have as one of its goals the aim of making redundant the presence in the community of academic linguists who are not from the community.

Czaykowska-Higgins, 2009, p. 25

Rice (2011) points out that in some parts of the world, communities already have guidelines for ethical research and/or research protocols in place to which outside field linguists must adhere. Canadian First Nations communities have been especially active in designing research protocols, see for example the discussion of the development of the Clearwater River Denesūliné protocol in Cheecham and Wilhelm (2013) and the Gwich'in protocol for traditional knowledge research (Gwich'in Social and Cultural Institute, 2004). In the United States, the Alaska Native Knowledge Network provides guidelines for respecting cultural knowledge and strengthening indigenous languages when working with Alaska Native language communities (Alaska Native Knowledge Network, 2000, 2001).

The discussion on ethics and linguist–community relations has also recently included viewpoints that are distinct from the collaborative research model. Most notably are Dobrin (2008) and Holton (2009), who point out that while collaboration toward linguistic goals is laudable in places like North America, Australia, and New Zealand, in other parts of the world like Melanesia and Indonesia, there are other standards for ethical behavior by outsiders. Communities may have nonlinguistic goals for their relationships with field linguists. Dobrin describes experiences with fieldwork in Papua New Guinea, where having a positive relationship with an outside researcher was perhaps more important than actual language survival, and where 'the right to language choice includes the right to choose against language' (Dobrin, 2008, p. 300). In the case of Holton (2009), differing expectations for his role as patron as well as linguist led to awkwardness in his relations with a speech community in eastern Indonesia.

Finally, there is an ethical concern for the long-term care of the data collected during fieldwork. Especially when so many of the world languages, and the knowledge systems encoded within them, are endangered, fieldworkers have a responsibility to make recordings that are of the highest possible quality, to describe languages accurately, to ensure the longevity of the data collected, and to disseminate them

responsibly. *Responsibly* in this case means finding a balance between the desires of the speakers for privacy and ownership of languages and recordings, and the need to manage one's data according to accepted best practices for fieldwork data sustainability. These best practices are discussed in the next section.

Data Management, Technology, and Archiving in Fieldwork

In their 2003 article, Bird and Simons brought to the forefront of the discipline of linguistics the importance of fully exploiting technology to create documentations of language that are long-lasting and multipurposeable. As field linguists increasingly took advantage of new technologies for capturing, storing, and disseminating language data, "new digital language resources of all kinds – lexicons, interlinear texts, grammars, language maps, field notes, recordings – [were] proving difficult to reuse and less portable than the conventional printed resources they replace" (Bird and Simons, 2003, p. 557). Since then, a great deal of research on digital processes, software, and infrastructure for linguistic fieldwork has taken place (e.g., EMELD, inNet). Technology and data management are primary concerns for the field linguist, who is responsible for not only documenting language but also seeing to the longevity and appropriate dissemination of the data.

The result is that field linguists must plan for and incorporate robust data management tasks into their fieldwork plans, and they must attend to these tasks both in the field while collecting data, and out, when undertaking postfieldwork processing. First, the linguist must plan for good-quality audio and video recording, including becoming familiar with the latest digital recording devices and microphones, understanding digital file formats like WAV and MP3, and learning how to arrange a recording environment for optimal sound or video quality with no background noises, disruptive light sources, etc. (see Margetts and Margetts, 2012).

The field linguist must also have a consistent plan for file management and metadata collection. File management tasks include selecting an appropriate file naming convention that gives each file a unique identifier, and developing a strategy for backing up files in the field. This may be especially challenging in field locations where electricity or dry storage conditions are in short supply. Metadata, which is key information about each file (e.g., date, language, location, demographic information about the people on the recording, etc.), must also be diligently collected (see Thieberger and Berez, 2012 for advice on data management in fieldwork).

The field linguist must also learn how to use and evaluate different pieces of software that she might use during data processing. For example, two of the most common software tasks for field linguistics are (1) time-aligning transcriptions and translation to audio and/or video, and (2) developing a lexical database. (*Time alignment* is the linking of bits of text (here, transcriptions and translations, but also possibly glosses, comments, notation of gesture, etc.) to the time code of a particular event in a piece of media. In many cases this is done by (a) referencing the media file and (b) referencing the

start and end times of, say, particular utterances within a marked-up text document of a transcription of that media file. See Berez (2007) and Thieberger and Berez (2012).) and (2) developing a lexical database. Before selecting software, field linguists must determine whether the output of the software is in an interoperable, human- and machine-readable format like XML, lest data eventually become inaccessible in the future, if the developer ever stops supporting the software. (Fortunately, there are already prominent pieces of software available for these linguistic tasks that field linguists use quite a bit: ELAN for time alignment of text to media and Fieldworks Language Explorer (FLEX) for building lexical databases and interlinearizing texts.)

A typical workflow for processing data collected during linguistic fieldwork is as follows. After the initial planning phases (which include determining a file name convention; coming up with a plan for backing up data in the field; and finding an archive to work with (more on this last below)), the fieldworker will make recordings during a session and also immediately collect metadata for that session. Audio and video data must then be transcribed, translated, and interlinearized, and lexical items must be entered into a lexical database. These items can then be used for analysis and in printed, online, and teaching materials, etc.

At each stage of the process – that is, after making recordings, after translating them, and after analyzing them – data should be regularly deposited in an archive. Field linguists should find a dedicated repository to work with before they leave for fieldwork, because the archive staff can make recommendations for proper depositing procedures. Many field linguists now make deposits via post directly from the field, or immediately upon returning home. There are a number of accepted digital language archives in the world, like those listed with the Open Languages Archive Community. Note that not every website, hard drive, or computer counts as an acceptable archive, and care should be taken that materials are safeguarded in a repository whose staff is knowledgeable about archiving language materials. Materials that are deemed private can be embargoed for a determined period of time, or access to materials can be made open to the public or restricted to a particular group of people (like members of the speech community).

See also: Corpus Linguistics; Fieldwork in Social and Cultural Anthropology; Language Policy; Language and Society; Linguistic Anthropology; Multilingualism; Pidgin and Creole Languages; Sapir-Whorf Hypothesis.

Bibliography

Alaska Native Knowledge Network, 2000. Guidelines for Respecting Cultural Knowledge. Assembly of Alaska Native Educators, Anchorage.

Alaska Native Knowledge Network, 2001. Guidelines for Strengthening Indigenous Languages. Assembly of Alaska Native Educators, Anchorage.

Austin, P.K., Sallabank, J. (Eds.), 2011. *The Cambridge Handbook of Endangered Languages*. Cambridge University Press, Cambridge.

Berez, A.L., 2007. Software review of EUDICO Linguistic Annotator (ELAN). *Language Documentation & Conservation* 1, 283–289.

Bird, S., Simons, G., 2003. Seven dimensions of portability for language documentation and description. *Language* 79, 557–582.

Blake, B.J., 1981. *Australian Aboriginal Languages*. Angus and Robertson, Sydney.

Bowern, C., 2008. *Linguistic Fieldwork: A Practical Guide*. Palgrave Macmillan, New York.

Bucholtz, M., 2011. *White Kids: Language, Race, and Styles of Youth Identity*. Cambridge University Press, Cambridge.

Cameron, D., Frazer, E., Harvey, P., Rampton, M.B.H., Richardson, K., 1992. *Researching Language: Issues of Power and Method*. Routledge, London and New York.

Capell, A., 1970. History of research in Australian and Tasmanian languages. In: Sebeok, T.A. (Ed.), *Current Trends in Linguistics, Linguistics in Oceania*, vol. 8. Mouton, The Hague, pp. 661–720.

Chafe, W., 1980. *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. Ablex, Norwood, New Jersey.

Cheecham, C., Wilhelm, A., 2013. *Sharing Worlds of Knowledge: Research Protocols for Communities*. <http://hdl.handle.net/10125/26110> (accessed 29.12.13.).

Chelliah, S.L., De Reuse, W.J., 2010. *Handbook of Descriptive Linguistic Fieldwork*. Springer, Dordrecht.

Crowley, T., 2007. *Field Linguistics: A Beginner's Guide*. Oxford University Press, Oxford.

Czaykowska-Higgins, E., 2009. Research models, community engagement, and linguistic fieldwork: reflections on working within Canadian indigenous communities. *Language Documentation & Conservation* 3, 15–50.

Dixon, R.M.W., 1980. *The Languages of Australia*. Cambridge University Press, Cambridge.

Dixon, R.M.W., Blake, B.J. (Eds.), 1979. *Handbook of Australian Languages*, vol. 1. Australian National University Press, Canberra.

Dobrin, L.M., 2008. From linguistic elicitation to eliciting the linguist: lessons in community empowerment from Melanesia. *Language* 84, 300–324.

Dorian, N., 2002. Commentary: broadening the rhetorical and descriptive horizons in endangered-language linguistics. *Journal of Linguistic Anthropology* 12, 134–140.

Dwyer, A.M., 2006. Ethics and practicalities of cooperative fieldwork and analysis. In: Gippert, J., Himmelmann, N.P., Mosel, U. (Eds.), *Essentials of Language Documentation*. Mouton de Gruyter, Berlin and New York, pp. 31–66.

Eckert, P., 1989. *Jocks and Burnouts: Social Categories and Identity in the High School*. Teachers College Press, New York.

Eckert, P., 2000. *Linguistic Variation as Social Practice: The Linguistic Construction of Identity in Belten*. Wiley Blackwell, Malden, MA.

Evans, N., 2009. *Dying Words: Endangered Languages and What They Have to Tell Us*. Wiley-Blackwell, Malden, MA.

Frawley, W., Hill, K., Munro, P., 2002. *Making Dictionaries: Preserving Indigenous Languages of the Americas*. University of California Press, Berkeley and Los Angeles.

Gerds, D., 1998. Beyond expertise: the role of the linguist in language revitalization programs. In: Ostler, N. (Ed.), *Endangered Languages: What Role for the Specialist? Proceedings of Second Foundation for Endangered Languages Conference*. Foundation for Endangered Languages, Bath, pp. 13–22.

Gil, D., 2001. Escaping Eurocentrism: fieldwork as a process of unlearning. In: Newman, P., Ratliff, M. (Eds.), *Linguistic Fieldwork*. Cambridge University Press, Cambridge, pp. 102–132.

Gippert, J., Himmelmann, N.P., Mosel, U. (Eds.), 2006. *Essentials of Language Documentation*. Mouton de Gruyter, Berlin and New York.

Grace, G. On the History of the Department of Linguistics at the University of Hawai'i at Mānoa, unpublished manuscript.

Grenoble, L.A., Furbee, N.L. (Eds.), 2010. *Language Documentation: Practice and Values*. John Benjamins, Amsterdam.

Grenoble, L.A., Whaley, L. (Eds.), 1998. *Endangered Languages: Current Issues and Future Prospects*. Cambridge University Press, Cambridge.

Grinevald, C., 1998. Language endangerment in South America: a programmatic approach. In: Grenoble, L.A., Whaley, L. (Eds.), *Endangered Languages: Current Issues and Future Prospects*. Cambridge University Press, Cambridge, pp. 124–160.

Grinevald, C., 2003. Speakers and documentation of endangered languages. In: Austin, P.K. (Ed.), *Language Documentation and Description*, vol. 1. School of Oriental and African Studies, London, pp. 71–101.

Gwich'in Social and Cultural Institute, 2004. Working with Gwich'in Traditional Knowledge in the Gwich'in Settlement Region. <http://www.grrb.nt.ca/pdf/GTCTKPolicy.pdf> (accessed 29.12.13.).

Hale, K., 2001. Ulwa (Southern Sumu): the beginnings of a language research project. In: Newman, P., Ratliff, M. (Eds.), *Linguistic Fieldwork*. Cambridge University Press, Cambridge, pp. 71–101.

Hale, K., Krauss, M., Watahomigie, L.J., Yamamoto, A.Y., Craig, C., Jeanne, L.M., England, N.C., 1992. Endangered languages. *Language* 68, 1–42.

- Harrison, K.D., 2007. *When Languages Die: The Extinction of the World's Languages and the Erosion of Human Knowledge*. Oxford University Press, Oxford.
- Hill, J., 2002. "Expert rhetorics" in advocacy for endangered languages: who is listening and what do they hear? *Journal of Linguistic Anthropology* 12, 119–133.
- Himmelmann, N.P., 1998. Documentary and descriptive linguistics. *Linguistics* 36, 161–195.
- Himmelmann, N.P., 2006. Language documentation: what is it good for? In: Gippert, J., Himmelmann, N.P., Mosel, U. (Eds.), *Essentials of Language Documentation*. Mouton de Gruyter, Berlin, pp. 1–30.
- Himmelmann, N.P., 2012. Linguistic data types and the interface between language documentation and description. *Language Documentation & Conservation* 6, 187–207.
- Holton, G., 2009. Relatively ethical: a comparison of linguistic research paradigms in Alaska and Indonesia. *Language Documentation & Conservation* 3, 161–175.
- Hyman, L., 2001. Fieldwork as a state of mind. In: Newman, P., Ratliff, M. (Eds.), *Linguistic Fieldwork*. Cambridge University Press, Cambridge, pp. 15–33.
- Ladefoged, P., 2003. *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*. Wiley-Blackwell, Malden, Massachusetts.
- Linguistic Society of America, 2009. Ethics Statement. http://www.linguisticsociety.org/files/Ethics_Statement.pdf (accessed 29.12.13.).
- Margetts, Anna, Margetts, Andrew, 2012. Audio and video recording techniques for linguistic research. In: Thieberger, N. (Ed.), *The Oxford Handbook of Linguistic Fieldwork*. Oxford University Press, Oxford, pp. 13–53.
- McGregor, W.B., 2004. *The Languages of the Kimberley, Western Australia*. RoutledgeCurzon, London.
- McGregor, W.B. (Ed.), 2008. *Encountering Aboriginal Languages: Studies in the History of Australian Linguistics*. Pacific Linguistics, Canberra.
- McLaughlin, F., Seydou Sall, T., 2001. The give and take of fieldwork: noun classes and other concerns in Fatick, Senegal. In: Newman, P., Ratliff, M. (Eds.), *Linguistic Fieldwork*. Cambridge University Press, Cambridge, pp. 189–210.
- Meyer, M., 1969. *Frog, Where Are You? Dial Books for Young Readers*, New York.
- Mosel, U., 2004. Dictionary making in endangered speech communities. In: Austin, P.K. (Ed.), *Language Documentation and Description*, vol. 2. School of Oriental and African Studies, London.
- Nettle, D., Romaine, S., 2000. *Vanishing Voices: The Extinction of the World's Languages*. Oxford University Press, Oxford.
- Newman, P., Ratliff, M. (Eds.), 2001. *Linguistic Fieldwork*. Cambridge University Press, Cambridge.
- Nichols, J., Sprouse, R.L., 2003. Documenting lexicons: Chechen and Ingush. In: Austin, P.K. (Ed.), *Language Documentation and Description*, vol. 1. School of Oriental and African Studies, London.
- O'Grady, G.N., Voegelin, C.F., Voegelin, F.M., 1966. Languages of the world: Indo-Pacific. *Anthropological Linguistics* 8, 1–197.
- Penfield, S.D., Serratos, A., Tucker, B.V., Flores, A., Harper, G., Hill Jr., J., Vasquez, N., 2008. Community collaborations: best practices for North American indigenous language documentation. *International Journal of the Sociology of Language* 191, 187–202.
- Rice, K., 2006. Ethical issues in linguistic fieldwork: an overview. *Journal of Academic Ethics* 4, 123–155.
- Rice, K., 2011. Documentary linguistics and community relations. *Language Documentation & Conservation* 5, 187–207.
- Rice, K., 2012. Ethical issues in linguistic fieldwork. In: Thieberger, N. (Ed.), *The Oxford Handbook of Linguistic Fieldwork*. Oxford University Press, Oxford, pp. 407–429.
- Rosenblum, D., Berez, A.L., 2010. Introduction: the Boasian tradition and contemporary practice in linguistic fieldwork in the Americas. In: Berez, A.L., Mulder, J., Rosenblum, D. (Eds.), *Fieldwork and Linguistic Analysis in Indigenous Languages of the Americas (Language Documentation & Conservation Special Publication 2)*. University of Hawai'i Press, Honolulu, pp. 1–8.
- Sakel, J., Everett, D.L., 2012. *Linguistic Fieldwork*. Cambridge University Press, Cambridge.
- Samarin, W.J., 1967. *Field Linguistics: A Guide to Linguist Field Work*. Holt, Rinehart and Winston, New York.
- Stebbins, T., 2003. Fighting Language Endangerment: Community Directed Research on Sm'algyax (Coast Tsimshian). *Endangered Languages of the Pacific Rim Publication Series A2–026*.
- Strand, L., Marullo, S., Cutforth, N., Stoecker, R., Donohue, P., 2003. *Community-Based Research and Higher Education: Principles and Practices*. Jossey-Bass, San Francisco.
- Strömquist, S., Verhoeven, L., 2004. Relating Events in Narrative. In: *Typological and Contextual Perspectives*, vol. 2. Lawrence Erlbaum Associates, Mahwah, New Jersey; London.
- Thieberger, N. (Ed.), 2012. *The Oxford Handbook of Linguistic Fieldwork*. Oxford University Press, Oxford.
- Thieberger, N., Berez, A.L., 2012. Linguistic data management. In: Thieberger, N. (Ed.), *The Oxford Handbook of Linguistic Fieldwork*. Oxford University Press, Oxford, pp. 90–118.
- Turner, K., 1997. Mary R. Haas: teacher. *Anthropological Linguistics* 39, 544–549.
- Woodbury, A.C., 2011. Language documentation. In: Austin, P., Sallabank, J. (Eds.), *The Cambridge Handbook of Endangered Languages*. Cambridge University Press, Cambridge, pp. 159–186.
- Wurm, S.A., 1972. *Languages of Australia and Tasmania*. Mouton, The Hague.
- Yamada, R., 2007. Collaborative linguistic fieldwork: practical application of the empowerment model. *Language Documentation and Conservation* 1, 257–282.

Relevant Websites

- <http://dobes.mpi.nl/> – Dokumentation Bedrohter Sprachen (DoBeS).
- <http://www.emeld.org> – EMELD.
- <http://tla.mpi.nl/tools/tla-tools/elan/> – ELAN Software for Time Alignment of Text to Media.
- <http://www.endangeredlanguagefund.org/> – Endangered Language Fund.
- <http://www.ogmios.org/> – Foundation for Endangered Languages.
- <http://fieldworks.sil.org/flex/> – Fieldwork Language Explorer Software.
- <http://www.hrelp.org/> – Hans Rausing Endangered Language Documentation Project (ELDP).
- <http://www.nfirc.hawaii.edu/lc/default.html> – Language Documentation & Conservation (Journal).
- <http://fieldmanuals.mpi.nl/> – L&C Field Manuals and Stimulus Materials, Max Planck Institute, Nijmegen.
- <http://www.innet-project.eu> – inNet.
- <http://www.language-archives.org/> – OLAC.