

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 fieldworkers⁴:

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

⁴See 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-elicite 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 Sųłiné, an Athabaskan 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 transcription 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

⁵For 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 language⁶ – 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

⁶Needless 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 inter-linear 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 is 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

⁸For 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/international 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 that¹¹:

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] *dalabborrd* ‘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>.

¹¹Translations 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.