

6 Field Linguistics

PAMELA MUNRO

1 What is “Field Linguistics”?¹

Unlike most of the other subfields of linguistics described in this book, field linguistics is not a theoretical discipline, and there is relatively little literature devoted to this area. Field linguistics, as I will use the term here, refers to the collection of primary linguistic data on the basic grammatical facts of a relatively little studied language in a relatively natural setting from ordinary speakers, and to the analysis and dissemination of such data. This type of data collection is usually called “fieldwork.” Classic fieldwork is done in “the field,” the area where speakers actually live (rather than in an artificial setting, such as a university classroom), or, even more classically, the area from which the speakers’ ancestors originated.

Many types of linguistic endeavor share some of these features of field linguistics:

- armchair linguistics, where a native speaker linguist reflects on his or her own judgments (often confirmed by questioning other speakers) and analyzes these;
- psycholinguistics, where speakers produce responses to highly controlled stimuli;
- language acquisition studies, where children’s language development is observed, often in a highly natural setting;
- sociolinguistics, where speakers’ linguistic behavior is observed and correlated with facts about their backgrounds.

Most people would agree, however, that these domains are not really field linguistics.

Although field linguistics can be done anywhere, it is not normally based on introspection: linguists working on introspective data usually are not field linguists, even if their language is quite exotic. Thus, the languages on which field linguistics is done typically have few if any native speaker linguists, and

one of the priorities of some field linguists is to train native speakers in the techniques of linguistic analysis.

A native speaker linguist might certainly use introspection to produce data to be analyzed for a basic description of his or her language, but introspective armchair linguistics is normally directed at puzzling out relatively obscure or at least higher-level problems in languages whose grammar is already fairly well understood. Similarly, psycholinguistic studies conducted in the laboratory, acquisition studies based on observation of children in their homes and elsewhere, and sociolinguistic studies conducted in a community generally do not have the goal of producing basic grammatical description.

Studies like these can succeed precisely because basic description already exists. The goal of field linguistics is to produce descriptions of languages – often the first such descriptions. For this reason, what I am calling field linguistics has also been called descriptive linguistics.

There are many techniques for collecting data and doing fieldwork (see section 2). But data collection is only the first step. The data collected must be analyzed (see section 4) and, very importantly, disseminated. (Data, even analyzed data, that remains in someone's notebook or computer or tapes is of little value to anyone.) Any circulated data must be written in a system that is analytically consistent and maximally useful to the widest range of users.

Although there is not much literature describing field linguistics as a field, the amount of linguistic literature that results from field linguistics is huge. The type of literature or other production that comes from the analysis of field data can vary considerably. Basic descriptions usually take the form of grammars (or articles on grammatical topics) or dictionaries. These works often serve as sources for reanalyzing the data, perhaps from a different theoretical viewpoint. Novel data from field linguistics has provided numerous vitally important insights to mainstream theoretical linguistics over the years, and may also be important for other scholarship (section 5). Many serious field linguists, however, feel a compulsion to make the results of their fieldwork available to the communities of speakers who use the language being analyzed (section 6). Because of these efforts, some field linguists may regard their work as having more social consciousness than many ivory tower enterprises, though possibly these feelings arise in part as a reaction against feelings that more theoretically oriented linguists hold those who collect primary data in low esteem.

Fieldwork is addictive, at least for some people. The reason I do field linguistics is that I feel energized and my spirits lift on days when I get to do fieldwork, and I cherish my relationships with the speakers I work with.

2 How is “Field” Data Gathered?

2.1 *Basic techniques of field linguistics*

Linguists gather data directly from native speakers of the languages under investigation. There are several ways in which this is done.

Most often, particularly in the early stages of fieldwork, a linguist uses an intermediary language in order to ask for translations of words, phrases, and sentences in the “target” language. This question and answer translation process is called “elicitation.”

Some field linguists frown on the process of direct elicitation and prefer to work entirely from more natural “volunteered” data. Most often, this involves recording from the speaker an extended narrative of some sort (a “text”), such as a retelling of a traditional story or a personal reminiscence. The linguist then works carefully through the text with the speaker, obtaining not only a careful transcription and translation but following up on grammatical constructions and paradigms that arise in the text, in order to put together a full description of the grammar of the language in the text. A counterpart to such text analysis, where possible, is observing natural conversations between speakers. Even if the linguist cannot understand everything that is being said, he can take note of new words and grammatical structures that may appear only in discourse.

Both techniques have their pluses and minuses. Beginning the study of a completely unfamiliar language with simple words in isolation is a good way to become familiar with the language’s sound system; hearing words only in complex context can make phonetic distinctions harder to hear than when those words are uttered in isolation.

But simple elicitation is never sufficient in itself. If the linguist makes up all utterances for translation or comments by the speaker, there is a significant possibility of creating unnatural or skewed data. For example, the speaker’s translations may be influenced by the structure of the intermediary language, or, when the linguist grows confident enough to create new forms and sentences on his own, the speaker may be too polite to reject these. (Consequently, it is important for the linguist to ask a speaker to repeat back any sentence he makes up himself – if the speaker cannot repeat it, it is unlikely to be fully acceptable – and to carefully mark in notes any sentence that was not spontaneously produced by a speaker. If a construction never occurs in spontaneous speech, but is only accepted on the linguist’s model, it is unlikely to be a standard feature of the language.)

Elicitation and textual analysis are important complements to each other. One cannot assume every grammatical structure will show up in a text, so it is important to elicit missing structures directly. On the other hand, texts and conversational data similarly may reveal words and structures that never appear in sentence elicitation.

Serendipitous events can produce spontaneous types of language that are hard to elicit and that may never appear in texts. I had studied the Muskogean language Chickasaw for eight years and hundreds of hours before I began bringing my new baby Alex to visit my Chickasaw teacher, Catherine Willmond. One day, she took him on her lap and patted with his hand on the table in front of them, telling him,

- (1) Pas pas pas aachi
pas pas pas say

I had never heard this type of sentence before, but discovered that it was a type of “expressive” construction used to describe noises that speakers feel is particularly appropriate for illustration presented to children. (Catherine’s remark could be translated either “He’s going *pas pas pas* (making a slapping noise),” or as a command addressed to him, “Go *pas pas pas* (make a slapping noise)!” The sentence was especially striking because outside of words used in this construction, such as the expressive syllable *pas*, Chickasaw has no words ending in *s*; other expressive syllables exhibit similar phonological peculiarities (Munro 1998). I have also learned that the presence of a baby is helpful for stimulating a speaker to produce diminutive forms of verbs, which in a number of languages may be used to show that a verb has a small or dear subject (somewhat like honorific forms in many Asian languages) (Munro 1988). Of course I am not suggesting that all linguists should bring babies into the field as a standard prop. But it is important to follow the speaker’s reactions and train of thought, and to pursue new lines of inquiry that are suggested by things that happen during the field session.

Up till now, I have not considered monolingual fieldwork, in which both the linguist and the speaker communicate only in the target language. Complete monolingual fieldwork is rather rare, since it requires enormous dedication by the linguist, if he is to really achieve a level of fluency such that he can discuss the speaker’s subtle judgments entirely in the target language. However, many other forms of grammatical study can be conducted monolingually, or partly monolingually. One considerable benefit of any such work is that it increases the native speaker’s respect for the linguist!

In what follows I will assume that field linguists will engage in some direct bilingual elicitation, but that this will be combined with other types of investigation.

2.2 *Getting started with fieldwork*

2.2.1 *The field methods class*

Many linguists’ first experience with working with a native speaker comes in a field methods course in graduate school. In such a class the students meet with a speaker of an unfamiliar language and elicit forms, which they transcribe and analyze. Eventually, the students learn enough to have a fairly good understanding of the grammar of the languages.

Part of field methods class involves learning what might be called politeness or respect. In certain stages of a field methods class, occasional students sometimes become so excited by the data that they forget that it is being provided by a real person, with a real person’s needs and feelings. (I’ve had students turn to me in the middle of a class elicitation session and say, “Why did he say that?”, referring to the speaker in the third person, as if he would not understand or be interested in hearing himself discussed!) The respect that is due to the native speaker who assists with a field methods class necessitates finding a

suitable word to refer to that speaker. Traditionally, the speakers who provide data for linguists are called “informants,” a word that originally had at least a neutral sense. In the last few decades, however (at least since Watergate), the English word *informant* has become a euphemism for *informer*, and it has acquired all the negative connotations of that word in the minds of most non-academics.² I see no reason to apply such a loaded, unpleasant word to the wonderful people who introduce me and my students to the joys of their languages, so I don’t use the word *informant*, and I don’t allow my students to do so in my hearing. Having to think of a substitute term is positive, since it forces the linguist – or field methods student – to evaluate his or her own relationship with the speaker. The normal term I use is “consultant,” but often (particularly when the speaker is older) “teacher” is more appropriate. Many of the native speakers who work with me are co-authors of books or papers about their languages; in this case, “collaborator” is probably the best term.

My own field methods classes follow a traditional model. The students are not told what the target language is until the first day of class, and after that they are asked not to read any literature on the language until they have figured out certain aspects of its grammar for themselves. I have them begin by eliciting nouns (since in most languages these can be pronounced in isolation more readily than other types of words); the class members discuss together first their initial phonetic transcriptions and then their first ideas about what the phonological system of the language is (what the phonemes or distinctive opposing speech sounds are, in other words). After the class members have worked out their own phonological analyses, we compare these to existing ones in the literature – if any exist – or attempt to work out a consensus, in the case of previously undescribed languages. (I discuss in section 4.2 below the question of how words in the language are to be spelled.)

I don’t allow students to tape record early class sessions in field methods. The reason for this prohibition is that no one initially is very good at recording data from a new language, however hard they try: only practice and analysis develop this skill. But if they know that a tape recording is available, many students are less motivated to work hard on transcription. In theory, having a tape of the session would mean that the student could work diligently on improving his transcription later. But a tape is never as good as being there with the speaker, when you can listen again, ask for repetitions, ask for slower or faster versions, or look at the speaker from different angles, so I don’t want students to adopt this crutch at the beginning. (We often do record a sample tape of interesting words after a few sessions, and students are welcome to tape sessions after they have learned to transcribe well, as long as they ask the speaker’s permission. It is wrong to tape record anyone without asking permission.)

Words in isolation are fairly easy for anyone to elicit from a speaker, but problems can arise when moving on to simple sentences. If I ask someone, “How do you say, ‘I’m going?’”, that person may tell me the way to say “I’m going” in his language, but he may also say “You’re going,” responding not to

Table 6.1 Verbal agreement markers in Garifuna

	(P)	(I)	(T)	(D)	(R)	(S)	(N)
1sg.	n-	-na-	-tina	-dina	-dina	-na	-nina
2sg.	b-	-bu-	-tibu	-dibu	-dibu	-bu	-nibu
3m.	l-	-ni-	-ti	-li	-i	-i	-ni
3f.	t-	-nu-	-tu	-ru	-u	-u	-nu
1pl.	wa-	-wa-	-tiwa	-diwa	-diwa	-wa	-niwa
2pl.	h-	-nü-	-türü	-dürü	-dürü	-ürü	-nürü
3pl.	ha-	-ya-	-tiyan	-diyan	-yan	-yan	-niyan

Source: Munro 1997

the metalinguistic translation task but treating the request like a real-world event. Students learn early that context is very important, since if the speaker imagines a different context from the one they have in mind, the result may be unexpected or confusing. Similarly, speakers learn how to interpret the strange questions linguists ask, and generally become much more tolerant of funny sentences. After students acquire a small vocabulary and learn something about the grammar of the language, they make up their own words and sentences, asking the speaker to judge if they sound all right. (This is a difficult skill for both the student and the speaker. Speakers sometimes feel it would be impolite to criticize an understandable but ungrammatical utterance by the linguist, while linguists in love with their own theories may not listen hard enough to the way the speaker says, "Yes, you can say that.")

In the second major assignment, the students have to work out how subjects and objects of different persons and numbers are marked in different types of clauses with different types of verbs. This assignment can be quite easy for some languages, or quite difficult, for others.

In American Indian languages, which have been the subject of most of my own fieldwork, and which I try to use as the target language for every graduate field methods course I teach, it is very common to find both subject and object marked with affixes on the verb or verb phrase. Sometimes such marking is quite transparent, but often it is not.

Table 6.1 presents the different verbal agreement markers in the Arawakan language Garifuna (spoken in Belize, Honduras, and neighboring regions of Central America). Markers in the P column are prefixes; those in the I column are infixes; and the remaining markers are suffixes. Each set of markers is distinguished for seven person-number categories: first person singular, second person singular, third person singular masculine, third person singular feminine, first person plural, second person plural, and third person plural. Although there is a considerable amount of overlap among the sets, they are all distinct.³

The markers in the different sets of Garifuna person markers are used in different syntactic environments: a member of one of the seven sets is used to

mark the subject and then, in certain constructions, a marker from a second set may be used to mark an object. Thus, for example, in a sentence like

- (2) N-áfaru ba-dibu “I will hit you”
 1s-hit aux-2s

two affixes are used: a P prefix *ba-* on the verb *áfaru* “hit” and a suffix *-dibu* on the future auxiliary *ba*. Speakers’ usage is entirely consistent, but understanding it is a challenge for the analyst, and requires a fairly extensive amount of data, since both the particular syntactic construction and the semantics of the specific verb involved determine which markers will appear. Because of the partial overlap between sets, only full paradigms can determine which set of markers is used in a given construction. Thus, for example, in (2), the *-dibu* suffix could belong to either the D set or the R set of markers. Only with further data, such as

- (3) N-áfaru ba-yan “I will hit them”
 1sP-hit aux-3pR

can we securely identify the suffixes in both (2) and (3) as belonging to the R set.

I have taught two field methods classes using Garifuna as a target language. Students have difficulty realizing the complexity of the pronominal agreement pattern (even taking into account the fact that they have not seen every marker in the data they are analyzing) – they are reluctant to believe that a system can be this complicated, and often prefer to assume that they may have misrecorded *-tibu* as *-dibu*, for instance, or to simply ignore troublesome pieces of data. The lesson here is to look for patterns and to accept that the data may be complex if that is the only consistent explanation.

An important class activity is analyzing a text from the speaker. After the text has been recorded on tape, students go through it individually, producing their own transcriptions of what they heard on the tape. Then we go through the text carefully with the speaker, as described earlier. Transcribing recorded texts like this in a language that one does not know well is extremely difficult. Although it is possible to produce a quick and dirty transcription of a recorded text by simply playing a bit of the text, asking the speaker to repeat what was on the tape, and writing this down, the effort of transcribing the text beforehand is worthwhile. Often speakers are mistaken about what was on the tape, or they may change an incomplete portion of the text to make it sound better out of context. Frequently, more than one version of the text is produced – a fully accurate transcription of the recording, containing pauses, hesitations, false starts, and so on, and the speaker’s edited version, with everything said right. Each of these has different linguistic uses.

From collecting data, students move on to more extended grammatical description and analysis, choosing individual topics on which to write substantive papers based on individual elicitation with the speaker. Many students’ field

methods papers are later revised for publication, or may even be developed into masters' theses or dissertations.

Increased student facility with and access to computation has changed field methods. In my current class, we exchange copies of all notes via e-mail, and have improved tape transcription with a sound editor. Word processing makes paper writing smoother, and the collected data can be searched in many ways in various data bases (see section 3.1.3). Still, there is no substitute for just sitting and staring at the data, as all serious analysts know.

2.2.2 *Finding a speaker*

The field methods class teacher locates a speaker and makes all the arrangements for that speaker to show up for class and elicitation appointments: students just need to come to class and use their brains. (Of course, this is an ideal situation: field methods consultants are people, not data machines, and they may get sick or develop other conflicts in the middle of the term, posing logistical problems for the teacher.) But real fieldwork requires the linguist to find a speaker to work with, which may be easier said than done.

One might assume that one would choose a language first, then find a speaker, and this is, of course, what many people do. But many other linguists who want to do fieldwork – but who, perhaps, are located in places where few exotic languages are spoken – happily choose to study any language that they can find a speaker of.

There are many ways to find a speaker. Personal contacts and serendipity are often very important. Because I know that every couple of years I will be teaching field methods, I keep up my contacts in the Los Angeles American Indian community. Los Angeles has a very large Indian population (largely relocated from reservations by now discontinued federal programs), but increasingly fewer speakers of Indian languages, and it sometimes takes me 50 or more phone calls to find someone. All Indian languages of the United States are endangered, most critically, so I know that eventually there will come a time when UCLA linguists will not be able to find speakers of more than a few American Indian languages in the city. But as long as immigration from Latin America continues, there will be a steady stream of speakers of indigenous languages from Mexico and further south. Many of these languages have never been described.

Many linguists have a lot of trouble explaining their theoretical interests even to members of their families. Before you look for a speaker to work with, it's important to consider how you will tell that person about your work and goals. I normally tell a speaker that I am interested in learning his or her language, and in my case (since I'm such a terrific language junkie), this is completely true. Field methods class presents a problem, however. I usually try to explain to prospective consultants that students take the class because (in our department) it is a requirement, and that they want to learn the process of learning a language from a speaker rather than from a language class or

from books or tapes. But this sounds a little cold, and it's not surprising that speakers have trouble believing that the students really may not be interested in their language for its own sake. I urge the students, therefore, to try to develop such an interest – to read more about the people and their culture than just about the language, and to work as hard as possible on their pronunciation. All of these help validate their interest to the speaker, and increase the speaker's trust. Doing these things, even if they start out as conscious behaviors designed to impress, increase the chances that the linguist will be successful, and really will learn a lot about the language.

3 What to Ask a Speaker, and What a Speaker Says

Some people begin fieldwork on a language with a definite question or agenda in mind. Perhaps they are researching a particular syntactic construction cross-linguistically, or maybe they are looking for data to compare with that in a related language they know better. Having too much of an agenda or coming to the work with too many assumptions, however, can produce unexpected results.

One linguist I know had an ambitious plan for a cross-linguistic study of the potential ambiguity of sentences with quantifiers, such as *Two men carried four boxes* (did they have a total of eight boxes, or only four boxes between them?). He had shown native speakers of a variety of languages cute pictures of various configurations of men and different types of boxes, with interesting results. When he asked the late Pollyanna Heath to describe the pictures in her language, Maricopa, however, he encountered problems. In Maricopa (as in many American Indian languages), verbs for various activities are selected based on the shape of affected objects. Since some of the boxes in the pictures were round and some were oblong, different verbs had to be used, and it was impossible to translate the sentences simply.

I was reminded that I didn't know everything about how to do fieldwork myself while I was studying Creek, a Muskogean language related to Chickasaw, which at the time I already knew very well. I was primarily eliciting Creek words to compare phonologically with those in other Muskogean languages, but also idly trying to learn a little about Creek grammar. After I had been working on Creek this way for about a year, I happened to ask my consultant, Betty Bland, for the translation of an English sentence containing a plural noun. I was chagrined to learn that Creek has noun plurals – I had never checked to find how these worked in Creek, because Chickasaw nouns have no plural form, and I wrongly assumed that Creek would share this feature.

For these reasons, it is good to begin work on a new language by doing a general survey of as many features of basic grammar as possible: verb and noun inflection, questions, negatives, existentials, passives (if they exist), causatives, reflexives, and so on. This procedure reduces the chance of embarrassing

surprises and often pays dividends in the form of revealing areas where the grammar is particularly worth studying.

If the language has been studied already, it is certainly worthwhile to review existing descriptions. These can be used to help develop a plan for early elicitation sessions, and may speed analysis. Of course, earlier descriptions may not be correct, or may prove to be based on a different dialect from that of the current speaker, so important facts from such works should always be rechecked. (This is not the only thing that should be rechecked, of course. The linguist's own data, particularly old data, should be rechecked and added to regularly. It's horribly embarrassing to find that a crucial word or beloved sentence elicited only once and cited frequently since then in fact turns out not to be replicable!)

As in a field methods class, it is best to begin the study of any new language with simple words in isolation in order to develop a feeling for the phonetics. Nouns are usually more simply inflected than verbs, so they are often good to start with. If a full sentence is too difficult to hear all at once, one can ask the speaker to say parts of it on their own.

Certain types of phrases, however, are dangerous to elicit out of context. I find that speakers of many languages are uncomfortable translating complex noun phrases on their own, and often translate "the blue house" as "The house is blue." To see how to say "the blue house," then, it is usually best to find out how this phrase appears in a sentence like "My friend lives in the blue house." (Actually, the same comment can apply to certain types of sentences. A complete sentence that may seem very easy to understand to you may be interpreted completely differently by the speaker. It is often useful to ask when a particular utterance would be used.)

It is always wise to note many things about elicited data. Obviously, if a speaker says a sentence made up by the linguist is bad, that is worth noting, but it's also important for the linguist to make sure that a sentence he makes up that the speaker approves can actually be repeated. A sentence that the speaker says "sounds okay," but which he can't repeat back, is certainly not a perfect sentence. Similarly, a construction which the speaker agrees to and repeats willingly, but which he never volunteers himself either in translation or in other uses, is an odd construction, and it's worthwhile for the linguist to try to figure out why this pattern is avoided in natural speech. Similarly, if a given sentence is translated by the speaker only with great difficulty, that should be noted too.

I try to always write down any comments the speaker makes about data we discuss. Catherine Willmond, my Chickasaw teacher and collaborator, occasionally says, "That's the way white people say it." This is a surprising comment, since I am the only non-Indian I have ever encountered who can speak Chickasaw at all, and the sentences in question are often completely novel for me. But evidently such sentences share some (incorrect!) feature with the speech of non-fluent speakers. I haven't figured this out yet, but I diligently note this comment each time it's made, along with other cryptic remarks. The late Robert

Martin, my first Mojave teacher, would explain the difference between two synonymous sentences by saying that one meant “You’re saying it” and the other meant “You’re telling him.” This is another one I haven’t figured out yet. But maybe some day I will!

3.1 *Working in the field*

3.1.1 Fieldwork can be done anywhere

I have made many field trips away from home to study languages. I spend an average of a week or ten days in Oklahoma (studying Chickasaw and Choctaw, and occasionally Creek-Seminole) every year, and I have worked on Yavapai, Mojave, Chemehuevi, Hopi, and Apache in Arizona; Zapotec in Oaxaca; Kawaiisu in California; Choctaw in Mississippi; and Alabama and Koasati in Texas, in each case on one or more trips away from home. But all the hours on all these trips put together would constitute only a small percentage of my total hours spent doing fieldwork on days when I spent the night in my own bed. Most of my field data has been gathered from native speakers with whom I met during classes at UCLA, in my office, or in their homes a few hours’ drive from Los Angeles.

Now, in a few cases the speakers in question were actually in their original (“aboriginal”) locations, since I’ve studied a number of California Indian languages that are still spoken less than half a day’s drive from where I live (such as Cahuilla, Tübatulabal, Luiseño, and Diegueño). Most of the speakers I’ve worked with in the Los Angeles area, however, are people who lived in Los Angeles, but who were born elsewhere. Most or all of my work with speakers of Zapotec (several languages, originally spoken in Oaxaca), Garifuna (originally spoken in Belize), Lakhota (originally spoken in South Dakota), Pima and Maricopa (originally spoken in Arizona), Navajo (originally spoken in Arizona and New Mexico), Cherokee (originally spoken in Oklahoma), Crow (originally spoken in Montana), and Yupik Eskimo (originally spoken in Alaska) was done in Los Angeles.

For the most part, linguistic data gathered away from speakers’ traditional homelands can be just as valid as linguistic data gathered in those homelands. But of course there are tradeoffs.

An important worry for many linguists contemplating working with a displaced speaker is whether that person still commands his or her language as well as someone with the support of a whole community. This is a valid concern – anyone can forget his or her language with no practice or stimulation. But any minority language speakers – as almost all speakers of American Indian languages are these days – are in danger of not using their language enough. Displaced speakers often use their language more than people back on the reservation – it all depends on their personal situation and circumstances. It is certainly important to chat with prospective consultants about how and

how much they use their language. And consultants may well bring different types of experience to different tasks. A field methods class, for instance, is primarily studying a single speaker's usage patterns – so it is not crucial that that speaker be a conservative follower of standard grammatical descriptions.

If the linguist contemplates writing the first description of the grammar of a language, it is important to work with more than one speaker, if possible, and to supplement work with displaced speakers with work in the homeland community. Even when the bulk of the work is done with a displaced speaker (such as Catherine Willmond, my Chickasaw collaborator, who has lived in Los Angeles since 1959), briefer exposure to other speakers can serve as a useful check on and addition to the data (for example, I have worked with over 40 other Chickasaw speakers in Oklahoma, some of them for over 20 years).

One really important and gratifying aspect of working with displaced speakers is that one can share the fieldwork experience with a much larger group of students and others than could ever come along on overnight excursions. I regularly bring Mrs Willmond and others to campus to introduce their languages to students who not only have never heard an American Indian language, but have never met an American Indian. Sure, we can tell such people about endangered languages – but meeting a speaker of such a language and experiencing first-hand the beautiful structures threatened with loss makes the point dramatically.

However, certain types of field linguistics can only be done where there are concentrations of speakers (as many as possible) located as near as possible to where their ancestors lived. Traditional dialect surveys are done only with the most conservative of speakers (never with those transplanted thousands of miles from home to a new multicultural environment); to be useful, these can only be done in the field. Other types of sociolinguistic data, particularly when relevant to a traditional cultural analysis, is also best gathered in a setting as nearly as possible approximating the ancestral one.

3.1.2 *Linguistics in the field is more than linguistics*

The main characteristic of actual fieldwork in the field – away from the ivory tower, specifically where the linguist does not get to sleep in his or her own bed – is that it's a 24-hour-a-day operation.

When I initially agreed to write this chapter, one of the editors of this volume opined that field linguistics must necessarily involve eating weird food and developing strange illnesses. Well, of course that is true. I have eaten grasshoppers (in Oaxaca) and squirrel (in Oklahoma); I have suffered from deeply embedded ticks that had to be surgically removed (in Oklahoma) and Montezuma's revenge (in Oaxaca)!

But weird food and illnesses are just part of the story, and not a very big part. What's different about fieldwork in the field is that the linguist participates in speakers' lives much more than when doing work with speakers in his or her own community.

A student of mine recently drafted a small grant proposal in which she estimated her daily mileage on a field trip at twice the distance between the motel she proposed to stay at and the location at which she hoped to meet with speakers. I suggested that this did not include the inevitable mileage spent driving around trying to meet speakers, or doing other things such as taking speakers without cars to forgotten doctors' appointments. Of course this isn't the fieldworker's job, but if you are there at someone's house with a car, won't you volunteer to take him or her to the clinic if there's no other way to go? Just as learning to be a good elicitor of data involves learning (or re-learning) basic politeness, learning to be a successful fieldworker means being willing to participate. It means not assuming that it's possible to make out a schedule of field sessions in advance (so many things intervene – especially, particularly when working with elderly consultants, funerals). And it means being ready to learn about other aspects of your consultants' culture. Being willing to give up your time to do this not only will prove to be personally rewarding, but will show speakers that you are really serious about learning their language. (An excellent memoir about linguistics in the field is R. M. W. Dixon's description of *Searching for Aboriginal Languages* in Australia (1984)).

3.1.3 *Technology and the fieldworker*

When I started doing fieldwork, there were no personal computers, and if I wanted to record a speaker I had to bring along a reel-to-reel tape recorder (and even the small portable models were bigger than a fat encyclopedia volume).

The first dictionary I did (a preliminary version of my Mojave dictionary) was compiled in three-inch by five-inch slips (some linguists, I know, prefer four-inch by six-inch slips!) – not cards (too thick!), but slips of ordinary paper, which were arranged alphabetically in a file box (one hundred slips take up only a little more than half an inch). Reluctantly, I have stopped introducing field methods classes to the joys of using file slips, which I still feel are unparalleled for their ability to be freely manipulated and arranged in different ways. But I don't use paper slips much myself any more, so it doesn't seem right to require students to make a slip file, as I once did.

Computers have changed fieldwork considerably, and they are now easily portable; with battery packs, they can be taken anywhere (and in fact solar chargers allow using them even where there is no electricity). With a portable computer, one can display and examine wave forms and pitch tracks, add to a growing database, and search for previous recordings and related data. There are now intricate programs for the construction of dictionaries and text analysis (though I still have found nothing that works as well for me as word processing programs).

Tape recorders have also improved exponentially in the last few decades. Even inexpensive portable tape recorders often produce excellent recordings, especially with a good microphone. The availability of high-quality digital recorders and microphones allows the recording of high-quality data, suitable

for all types of laboratory analysis, under the most difficult field conditions. Video tape recorders also allow any fieldworker to record gestures and other non-verbal cues, stimulating types of analysis hitherto never attempted with exotic languages.

4 Analyzing the Data, and What to Do with It

4.1 *Basic analysis*

The most useful way to find out what you do not know is to try to describe what you do know. It is very important to keep writing – sections of a grammar (or dissertation), papers, anything – and to try to see how well the language can be described within the framework of what you already know about language and how it works.

Linguistic analysis of many sorts is covered in other chapters of this book. The main point for a field linguist to remember is that analysis must be ongoing. The notion (which one sometimes hears) that a graduate student can go off to the field and collect data for a year, and then come back to the university and begin writing a dissertation seems ridiculous to me. The only way to know for sure what you need to know next is to have tried your best to understand and analyze what you have already learned.

The minimum sort of ongoing analysis, which I recommend to all my students, is to type up reports of each field session (or, alternatively, to enter new data in some sort of data base), preferably with notes, comments, and preliminary analysis. Looking critically at the data in this way helps to reveal gaps in paradigms and new directions to take in the next session.

4.2 *Writing the language*

An early goal in any sustained fieldwork should be to arrive at an understanding of the language's basic phonology. This is obviously easier with some languages than with others (though almost all languages present some tricky analytical issues). But without knowing which sounds are contrastive and what sort of allophonic variation may occur in which environments, the linguist is apt to get bogged down in low-level phonetic transcription and to miss significant generalizations.

The particular phonetic transcription system adopted is not too important (I think), as long as it is used consistently. My own colleagues who work on American Indian languages mainly use the "Americanist" symbols rather than the International Phonetic Alphabet (IPA) (for instance, *š* instead of *ʃ*), but neither of these has any particular advantage over the other, as long as one clearly sets out what particular symbols mean for the particular language under study.

Once the phonology is analyzed, it is worthwhile to adopt a clear phonemic orthography. Using a phonemic orthography simplifies the presentation of data and makes it easier to present one's analysis in almost any forum, except for certain types of phonetic or phonological study. Failing to use a phonemic orthography (if you yourself understand the phonemic analysis) is insisting on obfuscation: you are depriving the more casual reader of knowledge you possess. (One of the classic descriptive grammars of all time is Edward Sapir's description of Southern Paiute (1930–1). But few of the people who have praised this careful and indeed beautifully complete fieldwork-based study⁴ have spent much time with it, because it is exceptionally difficult for the casual reader, since it mixes at least three levels of transcription – very abstract phonemic, fairly superficial phonetic, and extremely detailed phonetic – and is often exasperatingly hard to work through.)

I strongly recommend that field linguists – and others working with languages that do not already have an established orthography – develop not just a phonemic orthography, but a practical orthography, one that can be written entirely on a standard keyboard (in other words, one that uses no special non-typeable phonetic symbols or diacritics). Using such an orthography means that one can enter data in any computer application (including e-mail!) without the use of special fonts, but it has a more important practical value. Ordinary people – native speakers and their relatives, scholars in other disciplines, and interested laypeople – can easily learn to read and use an orthography that doesn't make use of special symbols, but they are often mystified or even repulsed by an orthography that makes use of unfamiliar symbols. I have heard native speakers beg linguists to help them develop a way to write their languages without special symbols, but such pleas sometimes fall on deaf ears. This is odd, since the meanings of the symbols in a practical orthography can be explained just as clearly for the benefit of linguists (with a one-time use of IPA, perhaps) as other symbols can, so that everyone benefits.

Certainly, some languages are harder to devise orthographies for than others (particularly given the odd biases of current Euro-centered keyboards, which for example include *ã* and *õ*, but no comparable symbols for *e*, *i*, or *u*). But it is well worth it to put out the effort to develop such systems. (I discuss some of the problems of devising practical orthographies, and some clever solutions to these problems by a variety of field linguists, in Munro 1996, which incidentally presents an early orthography for San Lucas Quiaviní Zapotec, developed with my collaborator Felipe Lopez, that has now been modified several times!)

4.3 *Describing the language*

Some field linguists learn a lot about languages they work on, but never publish anything. This is a criminal shame, especially since the languages in question may not be spoken forever. I believe that any linguist who engages in extensive fieldwork has a duty to publish (or otherwise make available) as

much of his or her analysis of the language as possible. Preferably, such material should be disseminated in the form of clear description that is accessible to as wide a range of readers as possible. This is particularly true of languages that are seriously endangered, for which it is (alas) relatively easy to foresee a time when today's linguistic description will be the only source of information on the language.

The late Mary R. Haas, who founded the Survey of California Indian Languages at the University of California, Berkeley, and trained several generations of field linguists, taught her students that the most important goal of the descriptive linguist should be to produce a grammar, a dictionary, and a collection of texts. Such material can serve as the basis for production of pedagogical materials for language revival, cultural enrichment materials, background research in many disciplines other than linguistics, and later comparative and theoretical linguistic research. My own teacher Margaret Langdon wrote, "Only after seriously confronting (if not completing) such a task can one call oneself a linguist. On the other hand, I am convinced that this task cannot be approached without some theoretical assumptions to guide the enterprise and to provide the questions to be answered" (quoted in Hinton and Munro 1998: 1).

There are, of course, bad and good descriptions.⁵ A good description must be written with a solid understanding of the workings not just of the language being described, but also of language in general. For example, if someone making a dictionary has not worked out how many parts of speech the language has, with what morphological and syntactic characteristics, that dictionary will probably be incoherent.

It may come as a surprise to readers familiar only with European languages, but languages vary widely in just this regard. Although (I believe) all languages have verbs, nouns, and probably a few recalcitrant other types of words often called "particles,"⁶ many languages have no words corresponding to articles, and quite a large number of languages have no adjectives or quantifiers as we understand those terms with regard to English. (In many languages, adjectival notions are expressed by a subclass of either verbs or nouns, and I know quite a few languages in which quantifiers clearly are verbs, taking all expected verbal inflection.) But to accurately list and define words in a dictionary, the linguist must understand what the significant syntactic and morphological oppositions in the language are, and endeavor to encode these as clearly and accurately as possible.⁷

It is perhaps because of the field linguist's inevitable preoccupation with the minutiae of describing everything, of letting no piece of data escape unrecorded, that basic description is often dismissed as "pretheoretical". This term is sometimes used by theoretical linguists to mean that a description contains nothing relevant to current theory – no new constraints, no new projections. The irony is that such description is very often used as input to new theoretical advances (as I discuss further below), but it could not (or should not) be so used if it were not rigorously presented.

5 Contributions of Field Linguistics to Linguistic Theory and Other Scholarly Work

Basic descriptive data and analysis by field linguists contributes to the development of linguistic theory in two principal ways.

First, good description advances the theory by “testing” it, examining the way in which new data can be presented within current models, and showing how those claims must be extended and modified to handle new facts. Perhaps the most important early example of the importance of novel field data for the development of theory is Sapir’s seminal paper on “The psychological reality of phonemes” (1949 [1933]), which established the existence of native speakers’ mental concept of the phoneme (in Southern Paiute, Sarcee, and Nootka), foreshadowing the development of generative grammar. The best example I know of of a linguist who in his own work and that of his students has been constantly concerned with the relationship of field data to theory is Kenneth Hale, practically all of whose works present new and interesting data within a highly relevant theoretical context. Among the most significant is Hale’s work (based on languages of Australia and the Americas) on the notion of nonconfigurationality, which inspired extensive work on clause structure and pronominal and other arguments. Related work by Mark Baker, based in large part on fieldwork on Mohawk, resulted in important contributions to the theoretical treatment of incorporation (1988) and polysynthesis (1996), contributing to the development of the Minimalist program in syntax.⁸

Excellent contributions to linguistic theory based on solid fieldwork abound. Recent fieldwork-based dissertations by three of my students, for example, offered solutions to syntactic problems involving Binding in Choctaw (George A. Broadwell, 1990), Wh Movement in Western Apache (Brian C. Potter, 1997), and Antisymmetry in San Lucas Quiaviní Zapotec (Felicia A. Lee, 1999). Each of these works – like the best such descriptions – includes descriptive sections as well as theoretical argumentation.

Differences among languages provide valuable clues to how cognitive processes are related to speech, and the goal of much theoretical linguistics is to examine this relationship. But the theory can only be truly extended as it incorporates increasingly novel data-based observations. The relevance of these observations is not always immediately appreciated, so sometimes the most important contribution of a descriptive linguist will simply be to record facts about language that do not yet fit into any theoretical paradigm – but which will be relevant for future ones. Most typically, such pieces of data are noted by linguists without a theoretical axe to grind, whose whole purpose is to provide as complete a description as possible. Such people often note the existence of phenomena that are as yet irrelevant for current theory.

For instance, descriptive linguists have noted many ways in which pronominal agreement and case systems deviate from the Indo-European nominative-

accusative norm. Since the 1970s, these have become an important subject for typological analysis – Anderson’s (1976) and Dixon’s (1979) important studies of ergativity would have been impossible without a vast body of primary “pure” descriptions. More recently ergativity has been a concern even in highly theoretical work (e.g. by Hale and Keyser 1993 and Laka 1992).

Many other aspects of typological research advance through the work of much earlier descriptive linguists. When I was in graduate school it was a commonplace truism that no language had a basic word order that began with the object. SVO, SOV, VSO were accepted basic word orders, VOS had been observed in a few languages, but OSV and OVS did not occur – of course prompting the development of typological theories to account for this observation. Even as I was being taught about this, however, field linguists associated with the Summer Institute of Linguistics were recording Amazonian languages with just this word order. They did not describe these languages with an eye toward upsetting typological claims that they may not even have been aware of; they simply wanted to describe the languages they worked on thoroughly and well. But their work led to the advancement of typological studies (Derbyshire and Pullum 1981).

The study of phonology traditionally draws on a wider linguistic data base than syntax. Although Chomsky and Halle’s pioneering study of *The Sound Pattern of English* (1968) is now often viewed as the epitome of abstraction, this work set an important standard in terms of the number of languages that were cited in support of its claims (and, in particular, that went into the development of its feature system). In recent years, this trend has grown. Increasing numbers of phonetic studies have made possible sophisticated surveys of a very wide range of languages (e.g. Ladefoged and Maddieson 1996), which in turn provide input into theoretical studies of all aspects of phonology. Whole subfields of phonology, such as autosegmental phonology (1979), arose because of the realization that there were types of phonetic data that could not be handled easily within current theory.

Field linguists also contribute to other scholarly activity besides theoretical work in syntax, semantics, phonetics, and phonology. At the beginning of this chapter I mentioned several other areas of linguistics that rely on work with native speakers, such as acquisition studies and sociolinguistics. Such work cannot be done easily – or perhaps cannot be done at all – on languages for which no basic description exists, so providing basic descriptions lays the foundation for later linguistic analysis of almost any kind. Another field of linguistics for which basic description of as many languages as possible is vital is historical linguistics, and the related areas of classification and dialectology. Comparative and historical work must be based on basic field data.

Researchers in many other fields draw on primary linguistic description (and greatly appreciate it if it is as theoretically neutral and devoid of jargon as possible). Anthropologists, ethnologists, and historians make use of linguistic description for research both on contemporary populations and on historical records that may include material in minority languages. Scholars studying

place names, ethnobotany, and many other aspects of culture make use of primary linguistic description, particularly in the form of dictionaries.

6 The Highest Contribution

One of the most important reasons to do primary description is to preserve languages that may otherwise pass away. Languages reflect much of their speakers' culture and experience. Much of a society's knowledge and unique expressions will inevitably be lost with the disappearance of language. Linguistic diversity is one of the most visible and important aspects of mankind's cultural diversity, and language loss diminishes this diversity. The passing of any language reduces the range of human expressive power, and may lessen our chance of figuring out how language is realized in the mind.

Some linguists do extensive fieldwork in graduate school, write a dissertation, and then go on to careers as professional academics, publishing only on theoretical issues. Others (like me) never recover from the bite of the fieldwork bug, and must always go on to study one more language, or to learn one more word to add to the current dictionary.

The best contribution this last group of field linguists can make is to produce descriptions like those I have described here, which can be used not only as the basis of linguistic and other scholarly research, but also by the communities of the native speakers who have helped us, for assistance in language revitalization and cultural awareness programs or to promote literacy. Dictionaries that can be used by ordinary people, written with clearly explained, easily understood orthographies, and grammars (especially teaching grammars) that can be used by intelligent, motivated laypeople, are among the descriptive linguist's most useful publications.

These can also, of course, be the most enduring of contributions. Check the circulation records of any large library. With virtually no exceptions, the linguistic books that are still being borrowed 30 or 50 years after they were written are basic descriptions, not theoretical tomes.

NOTES

1 I am grateful to a number of colleagues who sent me their answers to this question and others I consider here: Aaron Broadwell, Ken Hale, Jack B. Martin, Laura Martin, Russell Schuh, and Siri Tuttle. I have learned a lot about fieldwork from observing and talking to many other linguists over the years.

I must also thank all the wonderful native speakers without whom I could not call myself a fieldworker, especially those I mention here: Betty Bland, Felipe Lopez, the late Pollyanna Heath, Catherine Willmond, and the late Robert Martin. As always this is for Allen and JP, and dear Alex.

- 2 There are especially unfortunate potential parallels between a traditional police informer and a linguistic “informer”: both are paid by an outsider in authority (surely a university professor is such a person) to reveal confidential information known only to the payee’s intimate circle. I believe that there are many potential non-financial benefits to a native speaker who works with a linguist – the work is often intellectually stimulating, the native speaker usually winds up learning interesting things about his language, and he may receive the gratification of contributing to his language’s preservation. Nonetheless, however, there are certainly groups who regard the teaching of their language to outsiders as a betrayal. Why should linguists use a term that invites this suggestion?
- 3 One might assume that the T set includes a morpheme *-ti* and the N set contains a morpheme *-ni*, which are added to the S set (with a rule deleting the first of two adjacent vowels), although it is difficult to suggest a meaning for these two morphemes. But this still leaves five separate sets!
- 4 Based for the most part, in fact, on data from a displaced speaker from Utah, Tony Tillohash, a student at Carlisle Indian School in Pennsylvania.
- 5 I could say a lot about bad description, but I won’t. There are good grammars and bad grammars, and good dictionaries and bad dictionaries. Usually it’s pretty easy to tell the difference just by inspection – inconsistencies and things that don’t make sense are pretty easy to spot if you look for them. But sometimes one can’t be sure one’s dealing with a bad description until one actually studies the language being described. This is scary, given that some languages can no longer in fact be studied! But the only solution is for more people to try to do the best job with description that they can.
- 6 Perhaps it is true that there are indeed languages for which there really is no distinction between nouns and verbs, but I have no personal experience with such languages.
- 7 A colleague once said (seriously, I believe, at the time) that you don’t even need to be a linguist to make a dictionary; all you have to do is write down words. This ignores the points just made in the text, as well as the need for a thorough phonological (and orthographic) analysis of the sort described earlier. I think that most likely this colleague no longer subscribes to this view, and perhaps spoke hastily even on this occasion. However, such remarks illustrate the relatively low standing of descriptive linguists in our field.
- 8 Polysynthetic languages are those that express many meanings within a single verb word. Baker’s definition of polysynthesis is more restricted than the usual understanding of this term, consequently (from my point of view) making the term less useful and interesting. But his claims based on his notion of polysynthesis are provocative and important.