

# First Steps in Ethno-Linguistic Fieldwork

Sebastian Drude and Hla Myat Thway

Freie Universität Berlin  
October 2003

## Preliminary Remarks

The purpose of the present paper is to provide an overview over important topics to be considered when it is planned to carry out linguistic fieldwork among a smaller local speech community with its own ethnical identity.

We are not trained specialists in this topic, but we hope that our own experience with field work could be useful to those who for the first time plan to do this kind of linguistic research. Nevertheless, the size of this paper is much too limited to give an introduction into linguistic fieldwork. (There are quite a few books and articles that focus on linguistic fieldwork, and the reader is referred to them.) Rather, this paper could be used as a checklist. Reflecting on each point we address could, as we hope, avoid some of the most common pitfalls and lessen some typical problems in the field situation.

We assume our reader to be an absolute layman in the topic. For anybody who has already conducted any actual field research, many of the points we raise would probably be obvious or trivial.

The present paper is intended mainly for linguists in Myanmar who are interested in conducting linguistic fieldwork among one of the smaller speech communities of Myanmar but have little or no experience yet. The presentation is organized into four larger sections. First we consider the relation of the researcher to the speech community (Sec. 1) and to the language informants (Sec. 2). Then we address the equipment and its use (Sec. 3), adopting a checklist with hints compiled by Gerd Klaas from the Max Planck Institute at Nijmegen. Finally we present one possible methodological research strategy (Sec. 4).

## 1 The researcher and his relationship to the speech community

We assume that choice of the language to be investigated has already been made, and that you know the basic facts about the speech community (SC) you plan to work with. Usually you will do an important part of your work in the local community where the language is spoken. This is necessary, among other things,

- in order to find the best speakers and language informants,
- in order to collect texts in a natural setting, and
- in order to get a basic understanding of the cultural and social environment for the language.

In some extreme cases it will be necessary to study a language outside its natural environment only, say, with some speakers who live temporarily or permanently outside the SC. But this should be taken as a major handicap for your investigation.

Before you do your first trip to the area, there are some tasks to be done. First, you should know as much as possible not only about the language but also about the living conditions, the culture and current situation of the SC. Has there been any ethnological research? Is there any report about the history of the people and about recent relevant events available? Try to get that information! You will need to be well-prepared to respect the local conditions.

Most importantly, is there anybody who has first hand information? Try to get into contact with somebody familiar with the local situation. This is not only important for you to get valuable, detailed, and more reliable information, but possibly it also gives you a good introduction into the group itself. A good point to start searching for such an introduction are people involved in, say, social organizations working in the area, e.g. in the health service or in an educational institution. Maybe there are former members of the local community that migrated to the capital or a nearby town? Try to get in contact with them!

Another important point is the overall political situation in the country or region. Are there different political groups or organizations that compete for power? How do they compete — are there armed conflicts? Is the SC affected or involved in such conflicts? The situation may vary a big deal from country to country, or even from village to village.

The general rule is that you should not get involved in any way in local or regional conflicts. Instead, you should try to establish neutral or good relations with all local forces (as good as possible to one group as long as this does not go at the expense of good relations to another group). Even if you need formal or informal research permissions by national or regional institutions or forces, you should take care not to be identified with any of the competing forces.

Once you get inside the village or local community (or, if possible, beforehand): You have to identify the local inner political situation. What are the relevant factions (parties, families, clans, etc.)? Who are the leaders, who are the followers, who seems to be or tries to be neutral? The rule is not to get involved in any conflict, and to try to establish good relations to relevant persons of as many different internal competing factions as possible, at least to more than one.

What does it mean to establish a „good relationship“ to the SC or parts of it? One of the first important point is reciprocity. You yourself know in advance what you want when you get there: do linguistic research that one day hopefully will lead to a good documentation and description of the language. But what does the SC want of you? Why should they possibly accept your presence and your work there? Try to find out in which way you could return the support of the SC.

Sometimes there are clear established traditional expectations on what somebody visiting the local community is supposed to do, such as bringing certain types of gifts right at the arrival. You should know of such customs and be prepared to attend them before you get there.

In many cases the form of retribution by the linguist to the SC can be established only during his stay. If the SC is poor, material help might be demanded. However, it will usually be the case that you by your own do not have the means to fulfill such wishes; in

this case, you should nevertheless try hard to get additional funding for such material help through appropriate channels.

If material help is not necessary or impossible at all for you to give, you should think of other ways of recompensation.

Consider if you can offer

- practical help of any kind
- to represent the cause of the SC in front of an institution or organization
- to contribute with your knowledge, e.g. by instructing local people, help in the local school, etc.

In any case you as a linguist should think of a way of returning the results of your research to the SC in a way that is useful to them, e.g. by preparing a local dictionary or didactical version of a grammar, or by developing a primer in the language, if the SC demands it.

If you manage to bring presents or gifts to the SC, be careful not to give preferential treatment to just one or a few factions. In many cases it will be the best to let the local authorities divide and distribute, insisting that every relevant part has a share in it.

On the other hand, you will have a closer relationship to some persons or subgroups, e.g., the family that accommodates you, a political leader that supports you, the families of the local teachers or of your language informants (these groups probably will partly overlap). If these subgroups do provide goods or services for you, they should be remunerated separately. Be sure to keep this apart from distributing general gifts for the SC as a whole.

An important point here is food. You should have enough foodstuffs with you to get along on your own (or enough money to buy it). Nevertheless, it is to be preferred to share your food with members of the SC (inviting them or giving food as a present) and to have the SC share their food with you.

Generally, your attitude should be one of

- *generosity*

If you have enough of valuable goods (medicine!), share them. If somebody asks for something you can do without, give it.

- *honesty and openness*

Never, never do a promise if you are not sure that you definitively will fulfil it. Explain the purposes and methods of your work to anybody who wants to know it.

- *respect*

Some aspects of the local culture may be obscure or even disconcerting to you. Try hard to accept and, if possible or necessary, adopt to local customs (without trying to go native!).

- *professionalism*

Of course there will and should be hours of leisure and you must relax sometimes from your work. But you are not there for holidays. You should never get involved in erotic affairs with local people, not even allusively or by insinuations!

- *patience*

Often things do not work out the way you hoped. Try again without getting nervous.

— *self-irony and humility*

You have to be aware that you WILL, sooner or later, commit mistakes. In such a case, apologise without being overbearing, don't get upset, join those who laugh at you (this is still better than the reaction of those who get angry).

It can happen (for reasons that are or are not in your responsibility) that the SC or relevant parts of it decide not to support your work among them any more and that they want you to go. In such a case, you should not insist stubbornly. Say that you are sorry, express your desire to return and to continue your work, and leave. If your field stay has been generally alright before, chances are that they will accept you back after a while or when the situation changes.

For an in-depth study of a language, you will need at least several (more than two) field trips of two to three months each. The ideal is to stay for half a year or longer among the SC and take the time to get some command of the language while doing the linguistic research. The absolute minimum are several short trips, each of a couple of weeks. In the intervals between the trips, you could work with one or a few informants in the city, if you have access to such language informants.

You should calculate that you will need between two and five years or even more to complete a thorough description and documentation of the language in form of a comprehensive grammar (i.e., a grammar including all of phonology, morphology, and syntax) and a dictionary.

## **2 The language informants**

You will do most of your work with the help of language informants (LIs). They will answer your questions, explain the meaning of words and sentences, repeat expressions slowly — in short, they will give you a better access to the language in question. So they are of key importance for your project.

The first important step is to choose the LIs you will work with for the most time, at least in the beginning of your work. (Possibly, other potential LIs will emerge later, maybe by showing constant interest in your work, see below.)

In each given situation the process of selection of the informants will be different. In many cases you will have no choice at the beginning, as LIs will be pointed at by local authorities or due to limited disposability or availability of suitable speakers.

There are several factors that potentially influence the choice of LIs.

— *Sex*

Usually, it is more relaxed to work with somebody of your own sex. If you adopt certain aspects of the local culture, you will have more shared experiences with a speaker of your own sex, identify the relevance of certain topics etc.

Work division in the local culture also will determine if women or men have more time at their disposal.

In many societies, it is the men that do the contact with the outer world and therefore are bilingual. So, at the first moment, you might prefer to work with men, as you can communicate with them in a common language. Later, when your own skills in

speaking the language eventually increase, you might prefer to work with the monolingual women.

Women tend to speak more carefully and be conservative as to language change, but they also might be shy, or intimidated by men who maybe expect them not to get involved in such type of work. (In rare cases, the opposite may also hold.)

Generally, the SC may have expectations on which kind of person is to do the work of a LI. As these expectations lead to approval or allowance of your choice, they are a factor of utmost importance, maybe even before and above the factual objective capacities of possible candidates.

— *Age*

You should choose speaker that have full command of the language in question. „Full command“ might be a culturally determined concept. In some SC speakers below a certain age (15, 18, 20, or even 30 or older) will be said to be not yet fully instructed in the language.

On the other hand, younger speakers tend to be more flexible in mind and can get trained more quickly as good informants. Physiological handicaps (e.g., less teeth) of older speakers can diminish the ability for certain tasks, too.

Age has also a great influence on disposability in terms of time. Middle-aged people with families might be too busy to dedicate constantly a lot of time in working with you, and therefore be less patient with your repetitive questions.

— *Education*

For most tasks, speakers with a formal education (bilingualism, literacy) and the capacity of abstract thinking are to be favoured. (In many cases, these will be local authorities who are engaged in education.) But education can also prevent a native speaker from taking a fresh, unbiased look at the language. (This is important in societies with their own traditions of reflecting on the language).

Education in terms of local culture is an asset for other tasks, e.g., taping of oral tradition or explanation of details of the local culture.

— *Social role and group*

The role of people in a society again has consequences for their disposability and the approval of your choice by the SC. In most cases, it might be wise not to choose a chief or factional leader (at least not as our only LI), due to possible political complications (within the group, or between you and your LI).

If you choose more than one LI, which will usually be the case, try to involve as many different political factions as possible.

Last but not least, personal sympathy and other individual factors should be mentioned. In any case, it should be made clear right from the start that the formal relationship of being a LI is a temporal one and that it may be necessary that this relation be ended one day.

Certainly, you can change or add LIs during your project, but keep in mind that all LIs — even those that eventually turn out to be of less use — tried to do a good job and deserve therefore your respect and gratitude. Ceasing to work with a LI is a delicate situation, especially if the remuneration is a major source of income for the respective family. Even mere adding a further LI to the inner circle can cause major embarrassment.

Therefore, you should not think of quickly replacing a LI thoughtlessly just because a minor problem emerged in the field. LIs tend to get better; the more you involve them into the theoretical side of the study of the language and its applications, the more their interest and capacities will grow. You should conceive training of the LIs as an important part of your work. Again: Be patient when things take a long time to work out ....

The relation between you and your LI should, again, be of reciprocity. He will have to spend hours and days and sometimes does hard and boring work. This must be compensated in some way, and this has to be clearly established beforehand. The best will be, in most cases, a determined and agreed payment or salary, but according to the given circumstances another form of compensation can be preferable or mandatory.

If ever possible, work with the LI alone, and usually only with one of them at a time. Besides better concentration on both sides, there might be persons that criticize the LI, maybe due to his speech varieties, claiming that he “does not speak right / well”. Such criticism can be based on good reasons, on envy, on political calculations, or on other reasons. In any case, you should prevent the informant to be exposed to such criticism.

Consider the personal rights of the informant, e.g., you should respect his wishes with respect to his voice or image. Do not give access to tapings, pictures or filmings to anybody without explicit prior agreement (possibly written) of the LI.

Never, never ever do “hidden” taping, that is, taping without prior explicit consent of the persons taped and without their knowledge!

### **3 Recording and processing: Equipment and taping basics**

#### ***3a General remarks***

You will not have a LI at your disposal wherever and whenever you would need one, and even the most patient LI would not want to repeat over and over the same word, sentence or text. So, you need to tape language data for later use and re-use.

At different stages of the research project, different types of tapings will be necessary. We will be concerned basically with audio tapings, which are absolutely mandatory to any language documentation and description project. Video tapings are an optional addition, necessary basically only for documenting the local culture and, possibly, certain physiological speech gestures (articulation movements).

First, you need audio tapings for a thorough phonological analysis of the sound system of the language. For this purpose, the most common means are „word lists“, that is, the expressions that correspond to (usually, are translations of) the content of listed stimulus words.

Later on you will tape sentences and texts. First you need shorter samples (tapings that range from a few sentences up to five minutes) for discovering basic grammatical features. This first step can overlap with later stages of the analysis of the sound system. Later you will need tapings of large portions of the local oral tradition or of (almost) natural language use.

Somebody who goes to the field for the first time certainly would like to know how much material he or she should tape. What would possibly be a minimum to be recommended?

There is no obvious answer to this question. It depends on several factors, the most important ones being manpower available and the frequency of (planned) field trips or the accessibility of LIs.

In the extreme case of having only one field trip to collect data that will be mainly analysed later, one could think that “the more, the better”, but be aware of the danger of being swallowed up by tons and tons of tapings that you cannot process afterwards. In order to give a first clue: One of the authors, who was in such a situation, collected about 400 small sentences (some with simple subordinated clauses), which also served to extract vocabulary items, and about 200 longer and more complex sentences. She also collected one coherent text, a story of about five minutes. This should do for a first survey.

The sentences to be elicited and taped should be carefully selected so as to cover many potential different constructions. In Bouquiaux and Thomas (1992), there are several word and sentence lists, including an extensive list for a language survey that could be adopted to your region. There is a general caveat: Even the most comprehensive corpus cannot tell you what is *not* possible to say in the language, and often this type of information is much more needed for your analysis than a hundred of correct sentences. Also, that a certain word or construction does not occur in even a very large corpus does not guarantee that the word does not exist, or that the construction is not possible. This means that no corpus can possibly substitute elicitation and interactive analysis with native speakers.

Besides sessions of taping planned beforehand (often in form of a monologue of a speaker in front of the recording equipment, performing a certain task, e.g., tell about a specific topic, give the equivalents to a stimulus word list etc.), you will have many sessions of interactive work with the LI, e.g., doing the transcriptions or translations or eliciting data in order to establish the grammaticality of a word form or a certain construction etc.

If you have enough storage capacity (tapes or digital media), you should also have a tape recorder running during most of these interactive working sessions with the LI. Listening to such tapings later can, in several ways, bring crucial but delicate points to your attention.

In what follows, we will summarize some important technical points in connection with tapings and the use of equipment. Most of it has been copied or adopted from a list prepared by Gerd Klaas of the Max Planck Institute (MPI) for Psycholinguistics in Nijmegen, Netherlands.

### ***3b Field equipment***

The MPI does not give specific recommendations on brands and types. Instead, there is a short list of equipment used by MPI people in the field. It does not cover all the available equipment, but it is a list of the most commonly used equipment.

Nowadays, modern audio and video recording devices are digital. Nevertheless, in many cases, you will have to get along with analogue devices, especially, with cassette (or tape) recorders.

You absolutely will need a recording device. Minimally a analogue audio recorder such as a cassette recorder (SONY WM-D6 analogue recorder) or at least a walkman. For documentation purposes (such as are indicated in the case of endangered languages), a digital device such as a Minidisk recorder is indicated.

You should use the best equipment available, as the quality of your recordings will be crucial for your later research, and for purposes of language documentation. The necessary equipment should always be sorted and checked beforehand and will cover the needs and the requirements for an expedition. Special things have to be requested in time. Test the equipment before buying it.

In any case, you should avoid dictaphones or other smaller devices designed for dictating (such as devices with smaller cassettes than the usual 10x6cm cassettes). Stereo is not necessary for most linguistic purposes, except you are focussing on dialogues or speech in cultural events with several participants.

There are good recording walkmans available. Larger desk devices are to be preferred. Be sure not to choose a device with automatic voice recognition (switching on only when the signal is strong enough). Such devices tend to cut off the first second(s) of the utterance and to hinder the quality. Also be sure to use a device with manually adjustable recording level.

Another important factor is the possibility to plug in an external microphone. An internal microphone is to be avoided as the mechanical noise of the recording device will diminish the quality of your recording. Also, the quality of an internal microphone tends to be bad in comparison to even a cheap external microphone.

Be aware that the quality of the microphone is the most important factor determining the quality of the recordings, above and before the quality of the recorder itself. So, a good external microphone should be used.

All devices can be supplied with the necessary supplies: Batteries, chargers, mains adapters, car adapters, solar operated buffered power supplies, 6/12 volt to 230V converters, microphones and special supplies: windshields, special cables for different mono/stereo microphones, radio-transmitter microphones, tripods of all sorts, special lenses, filters, carrying boxes, 'drying cells', etc. Most of this supplies are interchangeable, but some will only fit to special equipment.

Be sure to be familiarized with your equipment and to have all the necessary supplementary equipment with you.

In the case of a more sophisticated research project with capacities to acquire more expensive equipment, we give some hints about other recording devices.

For audio: You can use the DAT (Digital Audio Tape) recorder (SONY DTC D8 portable digital recorder, 44.1 and 48 kHz). Mostly recommended, however, is the MD (Mini Disk) recorder (SONY MZ-R90 portable digital recorder 44.1 kHz).

One of the authors has had bad experience with the DAT recorder, which proved too delicate for field conditions. In turn, the MiniDisk (MD) seems to be the most suitable modern recording device for audio recordings available. There should be always an analogue cassette recorder, at least as a backup solution, or for an additional second recording or playing for different purposes.

For video: use a DV (Digital Video) recorder (SONY DRC-TV990 and other, 3 CCD PAL digital video camcorder), or a HI 8 (with analogue eight millimetre tapes, such as the SONY CCD-TR XXX and many other analogue PAL camcorder)

If there is the possibility to use a laptop, we recommend a Tough notebook, a Sony VAIO, Siemens Fujitsu or other brands for showing audio, animations and video.

### **3c Recording tips**

Steps to be taken if you want to make a recording (audio only):

1. Before you go to the field: Get used to your 'field equipment'. Read the manuals! Test the complete set up before you have to use it in the field.
2. If the recording device has an indicator for the recording level, be sure it does fill in the green area most of the time, but does not pass over into the red area except for some very loud events.
3. Before starting the recording, you have to connect the microphone(s) to the recorder. The connectors of newer equipment are sometimes very tiny and sensitive for forces. So, look very carefully before you insert a plug into a connector.
4. If you use external microphones don't forget to switch them on. On most microphones a small light will indicate that the battery is in good condition.

For outdoor recordings some additional supplies might be useful:

- a) windshields for microphones
- b) some shields to protect the equipment from direct sunlight (equipment exposed to sunlight for a longer time may heat up so that device specifications are no longer valid)
5. The microphone should be placed in a distance of at least 20cm of the tape recorder. In order to avoid taping the mechanical noise of the recording device, you should use a damper such as a towel on the table where the tape recorder and the microphone will be placed.
6. The microphone is not to be hold in your hand or in the hand of the LI during the recording as any movement will cause noise being taped. Use a tripod or at least a small bag of beans or rice in order to fix the position of the microphone, which should be directed to the mouth of the LI.
7. The optimal distance between the microphone and the informant's mouth depends on several factors (generally, it should be between 30cm and 100cm). You have to adjust the recording level according to the sensibility of your microphone and the strength of the signal (volume of the speaker's voice).
8. Connect a headphone to the recorder to control the sound processing and quality.
9. Before you do your first planned taping, test the whole setup.
10. HUM, NOISE; can be caused by bad plug connections and / or the main power supply. Check all the cables and the wiring and / or use accus instead.
11. Do not leave the recording device in record / pause mode for a longer time. (In record / pause mode cameras uses more power and it's also a bad treatment for the tape. Some devices will switch off after 5 minutes).
12. A fly or a bee can disturb the sound processing heavily if turning around and around the microphone. Also windy situations can give serious problems. (Human ears may ignore these noises, a microphone never does!)
13. Generally, be careful to choose appropriate recording conditions. At least for purposes of phonetic analysis or language documentation, you should consider to do the recording outside the house or even outside the village, distant from children,

dogs, chicken etc. Another possible strategy is to do the taping later in the night when the majority of the SC is sleeping.

14. The first 5 to 15 seconds of any tape should never contain relevant data!! This is due to necessary 'pre-roll times' for editing and digitalization machines. Instead:
  - a) Make a short record (more than 15 seconds). Record your name, the date, recording place, name of the informant, an internal number for your organization and other relevant information via the microphone.
  - b) Don't erase this part of the tape. The information recorded will help to identify the tape even when the physical tape label gets lost.
  - c) Review and check the sound recording immediately. (In case of trouble check all cables, wiring, cassette protection, etc.).
  - d) If the sound is bad the recording heads of the tape recorder may be dirty. Use a cleaning tape for 10 seconds in play mode and don't rewind it. Alternatively, use a clean piece of cloths soaked in pure alcohol. If the picture/sound isn't better you can repeat the cleaning procedure once more. If the picture/sound is still bad, the device needs to be checked.
  - e) Directly after finishing recording, put the tape label a that tape and slide the tab on the cassette into 'Save' position to prevent from accidental erasure. In the case of ordinary audio cassettes, break off the plastic tongue on the back in order to prevent accidental re-taping.
15. In the case of audio recordings (cassette and tape), it is indicated to separate the different recordings on the tape by small pauses. This makes it easier to find certain parts for digitizing later if not the whole tape has to be digitized. DAT and MD recorders set ID marks at the beginning of a recording which might be retrieved for digitalization.
16. Do not use the entire length of a tape. (Begin and end may cause problems with digitalization process).
17. Never use the LP modus (long play) on your recording devices because it greatly reduces recording quality.
18. If the recording device switches the power off, because the accu is empty, take the accu off and charge it as soon as possible.
19. Keep an eye on your accus. Keep them charged.
20. If you want to copy the sound to an audiocassette, use the 'LINE IN' connector from the audio device even when the dubbing cable also fit's in the microphone input. MIC inputs are too sensitive for the strong line out signal of another device.
21. In a humid environment store your recording equipment in waterproof bags together with 'silicon gel'. The waterproof bag will also protect the equipment against dust. The silicon gel pack must be dried from time to time. You can dry them on a „barbecue“ fire, not in the fire. They are dry when the indication window of the cell show's „blue“. One can get these drying cells at nearly every expedition store.
22. Keep your tapes and recording material in a cool and dry place as long as possible.

23. Seal your recorded tapes with 'aluminium-foil' - or have them checked separately - when you pass the security check on airports.
24. If you use analogue audio cassettes or similar, you should use one side only, because the recording one side of the tape may spoil the recording on the other side.
25. Make backup copies of all important recordings as soon as possible. Use digitization wherever viable.
26. For a complete grammar you will need several hours, at absolute minimum half an hour of fully analysed texts and several hundreds of utterances of single sentences and word forms. A good corpus consists of several hours of transcribed and translated texts of different genres, capturing the most diverse speech styles (see below).

## **4 Working Procedures and Methods**

### ***4a General remarks***

Generally, there are as many different viable methodologies in linguistic fieldwork as there are good fieldworkers. Sooner or later, you will discover the best approach for your own style of work and for the language in question.

Few field manuals deal explicitly with working procedures, and we can also only give some starting hints and make some general remarks.

There are three main activities in linguistic fieldwork, and they complement one another and should go together: elicitation, work on recorded texts of (almost) natural speech of different genres, and analysis.

One more important activity is your own learning of the language studied, which will go hand in hand with your linguistic study, but will require extra dedication. You will have much more efficiency in eliciting or analyzing if you understand a great deal of the sentences you formulate in order to test hypotheses or that you get as an answer to a question or stimulus.

Several linguists with fieldwork experience insist that in the final stage of your work, all interaction with native speakers should be done in the native language. Whether you achieve this maximum goal or not depends on several factors. An important factor is the fluency of your LIs (and yourself) in the language used for interaction. The less the ease of communication, the more important is your ability in speaking the native language.

The same holds for the language you take your stimuli from (which might be identical with the language of interaction). If the LI has not a good command of the stimulus language, you will get poor results in elicitation and will have to double-check all semantic correspondences when your own command of the native language gets better.

In the following subsections we mention some of the important steps in the process of studying and documenting the language. You should keep in mind that these steps are not a strict sequence in time, but will be carried out with considerable overlap.

This overlap holds in particular for the two aspects of working with data: collecting data and analyzing data. If, in any area, you encounter questions which you cannot decide by examples in your corpus, you will have to elicit more specific data. This is why we cannot

give exact numbers for elicited forms, clauses and sentences. While working on the language you can expect to have several hundreds or even some thousands of elicited sentences of different degrees of complexity. Not all of them will be recorded or typed in digital databases.

The first three subsections focus on the three main types of data that will be most important on subsequent stages of your research: single words / word forms, small clauses and sentences of increasing complexity, and texts (mostly monological texts). The last subsection discusses the organisation of the data (on paper). Throughout this chapter we discuss the use of a computer where appropriate.

#### **4b The sound system**

The first step should be to establish the phonological system of the language (in case that nobody has already done so). Even if there exists already a writing system for the language in question, you should be sure about the phonological status of the existing conventions, so the best will be to start an analysis from scratch.

Pike (1947) wrote an early but still valuable „recipe“ to do such an analysis, following principles of traditional structuralism.

For doing a good phonetic / phonological analysis, you need a taped list of citation forms of different lexical words. You get these words by asking a speaker to translate stimuli lists. A usual word list would consist of about 400 words, optimally up to 600 words, a minimal list would of about 150 items.

For some regions in the world, there exist standardized stimuli lists that assure that your data will be comparable with those of other languages, especially for the later purpose of establishing the genetic relationship between languages. Find out if there is such a list for the region or language family in question (at least supposedly so), and use it for your basic „word list“.

That list is also the embryo of your lexicon. You will have to check and re-check later as there might be translation errors, misunderstandings, wrong or heterogeneous citation forms etc.

The next substep is to do a narrow phonetic transcription of these word forms, and to search for phonological contrasts. For reasons of consistency and in order to obtain comparable data valuable to other researchers, you should follow the conventions of the International Phonetic Alphabet (IPA). Be sure to have good command of these conventions and to take a copy of the most recent version of the IPA with you. The conventions are explained in the „Handbook of the International Phonetic Association“ (Cambridge University Press, 1999).

If you have a computer available, you could and should create a database. In early stages, it is problematic to use native words as the main sorting criterion, as the spelling is not established and might obscure more than help access to relevant data. You will probably use the numeric order of the stimulus lists or the equivalent forms in the stimulus language as headwords.

The program *Shoebox* is good at organizing such databases, because you can define different sorting orders for any language script, including transcriptions according to the IPA. You should in any case have a field in your database with a phonetic transcription in

each data set, and be careful to set up and adhere to your own conventions in typing phonetic data.

In some cases, your data may not be sufficient to establish or exclude a possible phonological contrast (e.g., a certain sound appears only in a given sound environment, and you cannot determine if this is by chance, or by a regularity). In these critical cases, you will have to complete your data with elicited data, e.g., by searching for word forms with a distinct meaning that differ only in respect to the critical sound (possible minimal pairs), or asking your LI if a given sequence of sounds possibly could occur in usual words etc.

You should double-check your transcriptions, try to read the forms to a speaker and pay maximum attention to possible corrections of your pronunciation.

By this, you may eventually come up with a set of discrete phonemes and their respective allophones and be able to determine their distribution (i.e. the conditions for their occurrence).

For a good phonetic description you should use computerized analyses of the sounds. An excellent program for phonetic analysis is *Praat*. It is freely available. Praat cannot take the decision for you whether a certain sound constitutes a phoneme or not. This is a question of analysis you will have to do. But you can find out much about the specific properties of certain sounds, and this can help you in your analysis. This holds especially for intonational features.

Once you have established a viable hypothesis about the segmental phonological system, you should create a working orthography for your own use that abstracts away from phonetic details. If you have a computer at your disposal, compatibility with computers and ease of typing is an important criterion for your orthography. Avoid diacritics wherever possible.

#### ***4c Forms and paradigms***

Later, or with temporal overlap, you should start to discover basic inflectional patterns of the words of different classes, so you will have to collect inflectional paradigms. Nouns and verbs of the elicited „word list“ are a good starting point.

You will need some few hundreds of elicited forms (maybe in small sample sentences focussing on those forms) to start, covering a wide range of types of verbs (transitive / intransitive etc.) and nouns that might constitute inflectional classes.

If there is any literature about related languages or languages of the same region available, you should study it carefully as it is probably the case that the language in question will share some or many of the grammatical features of those languages. But be careful not to anticipate the results, some crucial points may be different!

Also use your imagination to guess possible relevant grammatical form categories. Try to train the ability of the LI to give functional and formal equivalents. First, you will have to illustrate the abstract meaning evoking in detail possible utterance situations where the category in question would have to be used. You can't expect an LI to understand right from the start the technique of translating isolated inflected forms, but his capacity in abstracting from the concrete utterance situations (or of imagining these by himself) will get stronger with the time.

Even if you get a firm or certain answer, be sceptical about the status of the structure of the forms in the expression. In principle, you don't know if you are confronted with a single word, or an analytical word form, or other constructions, e.g., a combination of a word form and, say, an adverb that expresses the meaning of what is a grammatical category in the language of the stimulus.

So you have to check and counter-check with different forms or sentences, and with different informants. In early stages of the research, every answer has to be inspected, it might contain a valuable hint to a structure type in the language. Different answers might also be due to different language varieties (dialects, or styles, etc.).

Whatever your results, you may consider to organize sample forms in digital databases, too. This can be done in the entries of your lexical database (as samples with translations and indication of categories used) or can constitute a database type on its own. At some point you might want to set up a distinct database organized by grammatical content, for example, each category you recognized to exist in the language would have its entry, and the forms or sentences elicited could enter as examples here. Again, Shoebox is useful for such a work and comes with sample settings that you can adopt to your needs.

#### ***4d Work on texts***

When you have a certain basic knowledge of the sound system of the language and a first insight into the basic structures of word paradigms, you should begin to work with texts. Start with a shorter text, say, five to ten minutes. Ask a speaker to tell you about the origin of the local community, or ask him for a narration about some well-remembered historical fact.

Later, you should work with longer texts (ten minutes to several hours each) of different genres such as:

- narratives (mythological, historical, ...)
- explanations (focussing on central aspects of the local culture)
- reports, personal memories
- procedures (how to do certain tasks in the culture)
- dialogues
- ...

Transcribe the texts at least phonologically (using your practical phonological orthography). In the first weeks and months, you will need the help of a native speaker who repeats portions of the texts for you, speaking more slowly and well articulated.

In the case of the very first texts, you should ALSO do a narrow phonetic transcription. It is possible that you discover new sounds or contrasts not present in the elicited „word list“. Use elicitation to verify these new phonological facts, and re-elaborate your phonological analysis.

The computer can be a great help in this stage, too. Most modern computers will have a sound card and will thus be able to record and process sound events. You first „digitize“ the sound data, that is, you make a connection between your recording device (tape recorder, minidisk etc.) and the computer, play the recorded speech and enter it into

the computer, which will create a sound file, possibly in the Wave (.wav) format. There are several software programs that do this job, *X-gold*, *Cool Edit* among them.

A digitized sound file is much easier to handle when transcribing speech. You can select segments and play them repeatedly, which is much better than, e.g., rewinding a tape recorder. A program that is designed to do this and that can at the same time be used to type in your transcription is called *Transcriber*. It is freely available.

Besides a sound card, you need a device that reproduces the speech for you to listen. Many computers have built in sound speakers, but usually these are of a quality too poor for serious linguistic work, especially if you are not yet familiar with the sound system. In the city, where energy is easily available, you should consider to use external active loudspeakers (which will need their own energy supply, possibly from batteries). In a village, small headphones will be a good alternative. You will need Y-shaped adaptors that allow you to plug several headphones into the computer when you work with LIs.

After or while transcribing a text, you will have to translate it, also with the help of native speakers. Do both, a word-by-word translation and a free translation of segments of the text that you suspect to be sentences or clauses. (These hypotheses will be checked and possibly altered later.)

While doing the translation, start or continue elicitation in order to discover the grammatical structure of the sentences and phrases. Try to rearrange the order of the putative words, to let them out or to insert forms of adverbs etc. Discover, in this way, meaningful syntactic units above the word level. Use these word groups to expand your grammatical knowledge by varying, one by one, meaningful elements you suspect to be constituents on some level of grammatical organization.

At the same time, working on the texts will expand your knowledge of the lexicon. After a while, you should have a feeling of natural citation forms and forms that you could use to organize a lexical list.

Establish first hypotheses of grammatical lexical classes and subclasses. Use formal criteria such as structure of paradigms, and distributional criteria such as being or not being able to occur in a certain syntactic position. Keep in mind that such categories need not match from one language to another, e.g., what is an adjective in your language may correspond to a verb in the language in question.

#### ***4e Organisation of the data***

One major point is organisation of your data. You should take some hours every day, e.g., at night, to go through your annotations, organize the new facts and hypotheses, and formulate questions and topics to investigate for the next day.

We gave some hints on using the computer in organizing your data above, here we focus on handwritten material, which usually will not be replaced but complemented by the computer.

You will need several notebooks. Many researchers have one main „field book“ where they write down checked results, hypotheses etc., and several temporary notebooks they use to write results of elicitation and other data.

Usually it is best to use a pencil (if the paper gets wet, all brands of ink tend to get illegible) but NOT to use a rubber. Even corrected expressions or answered questions may

contain important hints for later stages of your study. So just cross them out, leaving a hint behind on where you wrote down the newer, valid information. Be sure to keep track of outdated data by adding the date to all annotations and organize everything according to the date.

For the analysis of texts, a classical arrangement is to use only the right side of the two pages in an open notebook for transcribing the text, leaving several lines free between the transcriptions in order to add word-for-word and free translations. The left side of the double pages can be used for questions, elicited information, or can contain references, including to pages in other notebooks where you elicited more information about a lexical unit or construction type.

Besides the notebooks for grammatical analysis, you will need to organise your lexical data. A traditional way are card files. They can be selected or rearranged according to your needs. This comes especially handy in the search for minimal pairs in the beginning of your research.

As mentioned above, an obvious choice is to organise much of your established (checked) data in a computer. There are small laptops that can be used under field conditions (see above). The very useful Program *Shoobox* (older versions are free) replaces the need for a card file for lexical data and can also be used to do analysis of texts.

## 5 Conclusion

We hope that these scarce remarks will be useful to new researchers doing ethno-linguistic fieldwork. Certainly we might have forgotten several points or even whole blocks of topics that would be important to mention. Other points will not apply in special conditions we have not thought of, and others simply may turn out to be wrong or inappropriate. We invite readers with more or other experience to improve on and add to this paper.

## 6 References and Suggested Readings

- Antworth, E. L. and J. R. Valentine (1998). Software for doing field linguistics. *Using Computers in Linguistics: A Practical Guide*. J. Lawler and H. A. Dry. London, Routledge: 170-196.
- Bouquiaux, Luc, and Thomas, Jacqueline M. C. 1992. *Studying and Describing Unwritten Languages*. Dallas, Texas: Summer Institute of Linguistics. (Firs edition 1976 in French.)
- Burling, R. (1984). *Learning a Field Language*. Ann Arbor, University of Michigan Press.
- Craig, C. (1997). Language contact and language degeneration. *Handbook of Sociolinguistics*. F. Coulmas. Oxford, Blackwell.
- Dixon, R. (1984). *Searching for Aboriginal Languages: Memoirs of a Field Worker*. Chicago, University of Chicago Press.
- Graves, T., Ed. (1994). *Intellectual Property Rights of Indigenous Peoples: A Sourcebook*. Oklahoma, Society for Applied Anthropology.

- Hale, K. (1964/65). "On the use of informants in field work." *Canadian Journal of Linguistics* **10**: 108-119.
- Harris, Z. S. and C. F. Voegelin (1953). "Eliciting in linguistics." *Southwestern Journal of Anthropology* **9**: 59-75.
- Kelly, J. and J. Local (1989). *Doing Phonology: Observing, Recording, Interpreting*. Manchester, Manchester University Press.
- Labov, W. (1972). "Some principals of linguistic methodolog." *Language in Society* **1**: 97-120.
- Labov, W. (1975). Empirical foundadions of linguistic theory. *The scope of American Linguistics*. R. Austerlitz. Lisse, Peter de Ridder Press: 77-134.
- Ladefoged, P. (1993). *A Course in Phonetics*. Fort Worth, Hartcourt Brace Jovanovich Publishers.
- Ladefoged, P. (1997). Instrumental Techniques for Linguistic Phonetic Field Work. *The Handbook of Phonetic Sciences*. W. Hardcastle and J. Laver. Oxford, Blackwell.
- Lee, R. (1995). *Dangerous Fieldwork*. Thousand Oaks, CA, Sage Publications.
- Longacre, R. (1964). *Grammar Discovery Procedures: A Field Manual*. The Hague, Mouton.
- Newman, P. (1992). "Fieldwork and Firdl methods in linguistics." *California Linguistic Notes* **23**(2): 2-8.
- Newman, P. and M. Ratliff, Eds. (2001). *Linguistic Fieldwork*. Cambridge, Cambridge University Press.
- Nida, E. (1947). "Field Techniques in Descriptive Linguistics." *International Journal of American Linguistics* **13**(3): 138-146.
- Nida, E. (1948). "A System for the Identification of Morphemes." *Language* **24**(4): 414-441.
- Payne, S. (1951). *The Art of Asking Questions*. Princeton, Princeton University Press.
- Pike, Kenneth. 1947. *Phonemics: A Technique for Reducing Langaue to Writing*. Ann Arbor: University of Michigan Press.
- Samarin, W. J. (1967). *Field Linguistics: A Guide to Linguistic Field Work*. New York, Holt, Rinehart, and Winston.
- Schütze, C. (1996). *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago, University of Chicago Press.
- Taylor, S. J. (1991). Leaving the Field: Research Relationships, and Responsibilities. *Experiencing Fieldwork*. Shaffir and Strebbins: 238-247.
- Vaux, B. and J. Cooper (1999). *Introduction to Linguistic Field Methods*. Munich, Lincom Europa.
- Voegelin, C. F. (1960). *Guide for Transcribing Unwritten Languages in Field Work*. Harvard University.