

Fieldwork on an Unfamiliar Language: Some Preliminary 'How To' Notes

One of the main skills this class seeks to develop is the simple skill of 'asking questions', a skill that has two principal components:

(1) The Two Principal Components of Skillful Fieldwork

- a. Manners
Asking questions in a way that keeps the speaker feeling comfortable (doesn't make them *uncomfortable*); generally keeping the whole experience a bearable one for the speaker
- b. Methodology
Asking questions in a way that yields information of maximal value to linguistic study.

Note:

These two 'sides' to the interview process are *not* in competition with one another. Indeed, one cannot hope to accomplish (1b) without continually maintaining (1a).

Although many of you are already very experienced in this kind of work, I thought I might share with you some of my own general notes on fieldwork methodology. These notes follow our general two-way division in (1).

1. On 'Manners'

"It may be something of a cliché, but all it takes is a little sensitivity and a willingness to learn." (Crowley, *Field Linguistics: A Beginners Guide*; p. 172)

While so-called 'fieldwork' these days can take a lot of forms, the one absolute, invariable, *sine qua non* is that you are working one-on-one with an actual person, a speaker of the language.

Thus, a good portion of the 'art' of the craft is 'interpersonal social intelligence', or just plain 'getting along with people and making them feel comfortable'.

While there are many little 'tips' that are good to keep in mind here, they all follow from the general, over-arching principle in (2):

(2) Respect

Treat the speaker with (friendly) respect:

Keep in mind that they are a human being, an essential 'collaborator' in the project, whose time is valuable, and who (particularly if they are older) probably know a good bit more about life and the world than you do.

From the general principle in (2) spring many smaller, general pieces of advice, which can also be organized according to various ‘subheadings’:

(3) **Friendliness**

- a. Always maintain friendliness and ‘cheerfulness’ during meetings.

(This can be difficult if you tend to be shy, in which case, you simply must do your best to overcome your inherent shyness/awkwardness.)

- b. Always maintain your own ‘engagement’ and ‘enthusiasm’ during meetings.

(Never be detached or perfunctory with the speaker.)

- c. Always be appreciative of the speaker’s contributions.

(i) *Greet all data appreciatively and interestedly, even if
-- it’s off the subject you want to talk about (after all, all data is useful!)
-- **it contradicts a ‘pet analysis’ of yours***

(ii) *Always thank the speaker for their time (even if you are also monetarily compensating them).*

- d. Never, in any way, be aggressive, challenging or confrontational with the speaker!

(i) ***Never get visibly upset (disappointed) if the data from the speaker challenge a pet hypothesis of yours!***

-- “Oh crap! I thought it would be X! Too bad, I guess!”

(ii) *Never challenge the data given by the speaker.*

-- “But, earlier you said X! That’s weird!”

(iii) *Never tell the speaker, in any fashion, that their judgment is ‘wrong’*

-- “But, the grammar says that the way you say it is X!”

“One of the first basic lessons in elicitation should be ‘Never tell your [speaker] that he/she has given the wrong answer.’ (Crowley, *Field Linguistics: A Beginners Guide*; p. 96)

(4) **Comfort**

Always strive to keep the speaker comfortable, strive to keep this from being a ‘negative experience’ for them.

- a. Keep things friendly! (see (3))
- b. Be careful not to overwork the speaker.
 - (i) *Always be mindful of the attention- and interest-level of the speaker.*
 - (ii) *Always be prepared to stop, take breaks, or move on to other things if they seem like they would appreciate it.*
 - (iii) *Don't heap question-upon-question on the speaker. Keep the 'intensity' of the interview nicely paced.*

(5) **Courteousness and Sensitivity**

Don't treat the speaker as a piece of furniture, or as some kind of ‘query machine’ that just ‘shuts off’ when you aren't talking to it.

- (i) Direct all questions about the language to the speaker
 - (ii) Look the speaker in the eyes when asking a question.
 - (iii) Avoid having extended (or heated!) arguments with team-mates in front of the speaker
-
-

2. On ‘Methodology’

The other, more specialized skill that ‘fieldwork’ requires is knowing *what kinds* of questions/tasks yield the *best kind of information* about the language.

Here, there are three very broad, fundamental points that apply across various sub-discussions.

(6) **Write Everything Down... *Everything!***

Everything, every little comment made by the speaker is valuable information.

Don't assume you can rely on the recordings or the transcriptions! Sometimes the recordings may be unclear, in which case your class notes will be of essential utility!

(7) **The Data Speakers Give Us Are Always (Merely) “Clues”**

The speakers of a language cannot on their own provide the answer to a linguistic question, even a question like “is X syntactically well-formed in the language?”

Rather, the information we get from speakers is a set of *clues*.

As the analysts (linguists), we use these *clues* to arrive at a ‘best-guess’ regarding the answer to the linguistic question we are interested in.

a. Example 1:

- We want to know if the syntax of the language allows ‘S(ubject)-O(bject)-V(erb)’ word order.
- We construct a sentence S that has SOV word order and ask the speaker “would you ever say something like S”. The speaker answers ‘no’.
- Has this conclusively shown that SOV is syntactically ill-formed in the language? NO!

The speaker has only rejected one sentence on one occasion. There could be a million reasons why they offered a negative judgment here.

- However, after we collect more data, it may be that – when combined with judgments like the above – the *preponderance of the evidence* suggests that SOV sentences are syntactically ill-formed.
- **Thus, the negative judgment alone doesn’t itself establish the answer to the question, but is ultimately just another ‘clue’ in the overall puzzle.**

b. Example 2:

- In the course of an interview session, the speaker comments “Oh, that suffix means that the verb is past-tense.”
- Has this conclusively shown that the suffix in question is past-tense? NO!

We don’t know anything about the empirical basis upon which the speaker is making this claim!

- Should we simply ignore this comment of the speaker’s? NO!

They might be on to something! Even the speaker’s own ‘pet theories’ about their language can stand as ‘clues’ pointing us to the proper analysis!

(8) **Never Ask (Avoid Asking) Questions that Require the Speaker to Do Analysis**

- a. *What is a question that 'requires the speaker to do analysis'?*
One that requires the speaker to offer or to judge the truth of a generalization.
- b. Examples of Such Questions:
- (i) Is that a noun?
 - (ii) When do you use that word?
 - (iii) Is that a focus particle?
 - (iv) When you say X, does it mean that the speaker already knows that Y?
- c. **Avoiding such questions can be very difficult, particularly if the consultant happens to be a linguist themselves, or someone that is generally knowledgeable about language (like a language teacher).**
- (i) A Real Example
When I was an undergrad, I did some work on Yoruba with a grad student in linguistics who was a speaker. At one point, I asked the student directly "Does Yoruba allow Strong Cross Over?" The student rightly shot back "You tell me!"
 - (ii) Another Real One:
A friend of mine once sat in on a field-methods class that was being taught by someone that wasn't the usual instructor. The consultant for the class was a visiting linguist, who had themselves done much work on the language. At one point the instructor (who was an excellent linguist) asked the consultant "Is that a focus particle?"
 - (iii) A More Common One:
You ask the consultant for the translation "I saw the bear". The consultant gives you a translation, but you notice that it doesn't contain a word corresponding to "the". As a follow-up, then, you ask the consultant "Is that sentence OK even if I wasn't already talking about a bear?"
- d. *What's wrong with these kinds of questions?*
- In order to answer the question, the speaker must **in their head** concoct other (similar) sentences/contexts, and then judge the naturalness of those sentences/contexts that they've constructed themselves. *But...*
- (i) The linguist should do the work of concocting those sentences/contexts!
 - (ii) Since the crucial examples/contexts are created in the speaker's head:
 - o You can't know whether the speaker is truly considering the *right* ones
 - o If they *are* the right ones, you are failing to record some key language data!

2.1 Major Methodological Issues and Controversies

Certain other points of ‘methodological advice’ are nicely introduced by considering some areas of continuing disagreement regarding ‘how fieldwork should be done’.

To illustrate these areas of disagreement, it helps to contrast two (rather insulting) archetypes: the ‘old-school’ fieldworker and the ‘old-school’ theoretician.

(9) Stereotype 1: ‘The Old-School Fieldworker’

- Works only at a ‘field site’, amongst a community of speakers.
- Collects massive amounts of ‘texts’ (recordings of fluent speech), which they then painstakingly transcribe with the help of the speakers.
- Never directly asks for ‘grammaticality judgments’ (e.g. “Can you say X?”)
- Uses the information gained from their texts (and only those texts) to create a dictionary and grammar for the language.
- The resulting grammar is about 80% phonology and morphology, 19% syntax (consisting of very superficial statements), and 1% semantics (mainly just consisting of rough ‘glosses’ for various items)
- Prides themselves on not knowing anything about linguist theory, or letting ideas in linguistic theory influence their work. (Says things like “We shouldn’t force a language into our preconceived, Euro-centric notions of language”.)

(10) Stereotype 2: ‘The Old-School Theorician’

- Works only in their office, on their own native language (usually English)
- Considers only their own grammaticality judgments (e.g. “Can I say X?”)
- Never looks to natural language texts or corpora for their data
- Uses the information gained from their own grammaticality judgments to create a ‘paper’, typically criticizing some earlier analysis (of English) and arguing for a novel one.
- Prides themselves on letting only linguistic theory guide their work, rather than chasing after data that look ‘exotic’. (Criticizes those who do otherwise as “butterfly collectors”)

While very few individuals have ever hewn very closely to these extremes, it’s long been recognized that neither extreme is the most productive way to proceed in linguistic study.

(11) **Weaknesses of the ‘Old-School’ Theoretician’s Way of Doing Things**

The data obtained are incredibly narrow (and perhaps ‘problematic’):

- a. Reflect only languages spoken by professional linguists (typically English)
- b. Reflect only a single type of data: judgments regarding the ‘naturalness’ of sentences denuded of any context (which can sometimes be misleading)

(12) **Weaknesses of the ‘Old-School’ Fieldworker’s Way of Doing Things**

- a. The injunction that one should *only* do linguistic research when living within the language community *severely limits the number of people that can do linguistic research on a language other than the one they speak natively.*

Look, some of us have families and can’t just jot off to Kenya whenever we want. But, sometimes we’re lucky and run into someone in our home country that speaks (*e.g.*) Dholuo.

Why deny ourselves (and the field) the opportunity to study this language?

- b. From a certain perspective, texts are actually a rather *poor* source of information. (more below)
- c. Willful ignorance of linguistic theory (particularly nowadays) only creates barriers to one’s understanding of the language (more below)

(13) **On The Poverty of ‘Texts’**

A text can tell the linguist a great deal about the structures the language *allows*.

But a text often says nothing about the structures the language *disallows*.

- a. Just because a structure doesn’t appear in a text (or collection of texts), it doesn’t conclusively follow that the structure is *disallowed*.
- b. However, if we were to directly ask a speaker for a *grammaticality judgment*, that *could* provide some stronger (though still not conclusive) evidence for ill-formedness.

The issues in (13) are especially relevant to the study/documentation of the *syntax* and *semantics* of a language, which typically require that linguists examine complex structures that one is not likely to ever encounter in a text.

(*e.g.* How do we tell if some phrase got to a position X by ‘movement’?)

(*e.g.* How do we tell if a certain phrase is ‘definite’ or ‘indefinite’?)

(14) **On the Value of Linguistic Theory to Fieldwork**

Consider the following claim, which is a caricature of a general perspective that one does still sometimes encounter:

a. Linguistic Description as ‘Cramming Languages into a Pre-Determined Mold’

“We don’t want our preconceptions about language to color our analyses of particular languages. That would just be like trying to force the language into a pre-determined ‘mold’ set by Euro-centric ways of thinking. Indeed, probably the best, most objective linguistic study would be done by someone who has never had a class in linguistics, and has no preconceptions of what languages look like!”

Majority opinion in the field is that such an extreme view is ‘not a productive way to proceed’. *Knowledge of linguistic theory is always useful, as long as one tempers that knowledge with skepticism!* Here’s why:

b. At this point in its development (*ca.* 2010), linguistic theory has been informed by the study of a great variety of human languages, around the world and throughout history.

- *Thus, even if all currently existing theoretical proposals are ultimately bogus, by studying them one does learn a great deal about what the languages of the world are like (what kinds of linguistic variety you see in the world).*

c. The more you know about what kinds of things human languages do, the better prepared you are to recognize ‘familiar patterns’ in the language of study.

- *If you’ve never seen or heard of (say) an ‘infix’ before, then think of the headaches that Tagalog would give you!*

d. Regarding the concern over ‘forcing a language into a mold’, *if the glove fits, you must convict!*

- If the (‘Euro-centric’) analysis is indeed ultimately wrong, then further study will show that it’s wrong.
- *We should always be engaged with finding the best possible analysis for a given set of data, and if that analysis happens to be one that works for a European language, so be it!*

e. Linguistic theory is an impetus to further language documentation!

- Without the theory of ‘movement’, why would anyone spend much time studying/documenting long-distance dependencies in a language?

(Indeed, prior to the theory of movement, they didn’t!)

While there are the above issues with the methodology of ‘the old-school fieldworker’, there are two points on which it is absolutely correct:

- The value of collecting and transcribing ‘texts’
- The methodological problems with collecting ‘grammaticality judgments’

(15) The Value of Texts

A ‘text’ is a document of fluent, connected, natural discourse in the language. As such, it can expose the linguist to grammatical structures that they wouldn’t otherwise have encountered.

- a. Sometimes, in one-on-one translation sessions, speakers avoid using grammatical structures that are ‘complex’ or ‘artful’ or otherwise ‘marked’, and instead use structures that are (in a general sense) ‘simpler’.

However, in natural fluent texts, these more complex structures shine!

A Real-Life Example:

A friend of mine once had been studying the relative clauses of a language for months before he started collecting natural texts in the language. Almost immediately, he found that the language possessed *two different* strategies for relativization, one of which he had never seen before!

- b. Since they are connected discourses, texts in some languages contain markers signifying ‘discourse-related’ properties like ‘topic-hood’. Such markers are often omitted in the single, isolated sentences speakers are asked to translate during interviews.

For these reasons, it’s widely accepted that the collection of ‘texts’ is a crucial part of the overall field study of a language. We will, later in the semester, engage in the collection and transcription of such ‘texts’...

(16) The Problems With Asking for Grammaticality Judgments

It’s actually still somewhat controversial in some circles whether linguists should *ever* ask speakers for ‘grammaticality judgments’, that is, to ask speakers to judge whether forms constructed by the linguist are ‘natural/acceptable’ or not.

“The good interrogator restricts himself to asking questions – he doesn’t make up novel utterances in the language of his informant just to see whether the informant will accept has he has composed.” (Harris & Voegelin 1953)

The general reason for this, explained below, is that *speakers can accept or reject a structure for a variety of reasons, many of which are not necessarily linguistic.*

(17) **Reasons Why a Speaker Might Accept a Structure Composed by the Linguist**

- a. The structure is well-formed and pragmatically natural in the language.
- b. **Although there are ‘problems’ with the structure (it’s not truly well-formed or pragmatically natural), the speaker is *giving the linguist a break*. (because *e.g.* the linguist is only a beginner in the language)**
- c. Although there are problems with the structure, the speaker is tired and doesn’t care anymore.

(18) **Reasons Why a Speaker Might Reject a Structure Composed by the Linguist**

- a. The structure is either ill-formed, or is not pragmatically natural.
- b. **Although the structure is well-formed, the speaker is not imagining the right context for its use.**
- c. Although the structure itself is perfectly fine, the speaker is *being hard on the linguist*. (because *e.g.* the linguist has failed to pronounce certain words correctly)
- d. Although the structure is in fact perfectly fine, the speaker is tired and doesn’t care anymore.

(19) **The Upshot**

While the ‘confounds’ listed in (17) and (18) are indeed potential issues, they are not necessarily insuperable or fatal methodological flaws. *The linguist must simply be careful to:*

- a. Control for the ‘confounds’ in (17b,c) and (18b-d).
 - For (17b), make sure to ask the speaker “would *you* ever say X” rather than “would it be OK for *me* to say X”?
 - For (18b), make sure to spell out for the speaker a context that (you think) would legitimize the structure.
- b. Ask for judgments on a variety of sentences (+ contexts) exhibiting the structure the linguist is interested in. Ask a variety of speakers on a variety of occasions.
- c. As with all data, remain skeptical, and be prepared to revise your views! (Remember the ‘slogan’ in (7): these data are ‘clues’, not answers!)

2.2 Some Notes on Asking for Translations

Even more than ‘asking for grammaticality judgments’, a major part of our ‘toolkit’ will be asking the speaker to translate structures of English into their language.

(20) When to Ask for Translations?

- a. When you have no idea how a given proposition would be expressed.
(*What do relative clauses in the language even look like?*)
- b. **When you want to know whether a structure is acceptable or not.**
 - (i) The Procedure:
 - Start with a sentence that you *know* to be acceptable, one that the speaker has themselves given to you.
 - Change that sentence *minimally* to test whether the targeted structure is acceptable or not.
 - (ii) Example: (Is SOV Order Possible?)
 - Assuming we know the words for “dog” “man” and “chase”, ask for the translation of “The dog chased the man”.
 - Ask whether *the speaker* would ever use the same set of words *in SOV order* as a translation.

(21) What *Kinds* of Things Can We Ask For Translations Of?

- a. Sentences (Well-formed sentences of English)
- b. So-called ‘open class’ or ‘content words’ (e.g. “jump”, “table”, “dog”, *etc.*)
- c. Avoid Asking for Direct Translations of Anything Else
 - (i) So-called ‘closed class’ or ‘function words’ (e.g. “all”, “the”, “every”)
*These often can’t be translated into other languages.
Their meaning is often hard to intuitively identify.*
 - (ii) Anything below the level of the word (e.g. “-ed”, “-s”)
*Tons of problems with trying to do this...
At the very least, it has all the problems of asking about ‘function words’.*

Caveat: Even asking for single open-class words can be problematic.

- (a) The language may lack the word in question (*cf.* English “go”)
- (b) English is inflectionally poor, and the language may ‘translate’ the English word differently in different syntactic contexts (e.g. what is the Italian translation of “eat”?)

(22) **Could the Speaker's Translation be Influenced by English?**

Especially in older work, one sometimes encounters the following concern/criticism:

a. The Influence of English on the Translation

In translating an English sentence into their language, a speaker may be 'influenced' by the grammar of the English sentence, and produce a translation that is not entirely natural in the language itself.

However, in response to this worry, it's fair to say that many (most?) in the field adopt the following set of views:

b. If a given structure in English is fully *impossible* in the language of study, then it's *highly unlikely* a speaker will ever use that structure in their translation.

(i) Thought experiment:

How would *you* translate the Spanish sentence "te amo"?

Would you ever feel inclined to translate it as "you love" or "I you love" or "you I love" or "you love I"?

c. If a given structure in an English sentence is possible in the language of study, then it *may* sometimes happen that a speaker will use that structure in their translation, even if it would (for whatever reason) not be as natural a sentence in the language of study. HOWEVER:

(i) Subsequent investigation will eventually reveal that the structure in question isn't the *only* way to express the given proposition in the language of study.

(ii) If we are at all interested in the most *natural* way to express a given proposition, then we will certainly turn to 'texts', where the likelihood of such 'emulation of English' is essentially zero.

3. **Some Additional, General Points of Advice on 'Fieldwork' (Broadly Construed)**

(23) **Try to Learn the Language You're Studying!**

Not only is it fun, it greatly aids your own work in the language.

At the very least, try your absolute best to pronounce and comprehend the sounds!

"[I have worked on] as many as eighteen different languages, though I never tried to carry out even a basic conversation in most of these." (Crowley, *Field Linguistics: A Beginners Guide*; p. 155)

(24) **Don't be Afraid to be Wrong!**

No matter how much we learn about a language, some part of our overall analysis is going to be *wrong!* (e.g. English, for crying out loud!)

This knowledge that we are all doubtless going to be wrong about *something* should be liberating!

- a. *Never be afraid to (provisionally) leap to an analysis or hypothesis, even though you've only looked at a small set of data!*
- b. *On the other hand, always be ready to revise your assumptions – any assumptions! – about the language, on the basis of further data.*

Example: The Existence of a 'VP'

- From our knowledge of syntactic theory, we may be used to the idea of dividing a sentence up into a 'Subject' and a 'VP'.
- However, in English, this analysis (which we all take for granted) is actually based upon a variety of *empirical* arguments (VP fronting, VP ellipsis, Subject-Verb agreement, *etc.*)
- Should we, then, avoid imputing a similar 'VP projection' to our language of study, until we can rustle up some similar empirical arguments?
- Some would say *YES!* But, here are some reasons to say *NO!*
 - (i) In order to do any analysis at all, we're going to have to provisionally assume *something* about the clausal structure of the language.
 - (ii) As long as we're clear in our notes and our writing that the 'VP analysis' is a pure stipulation, we needn't risk confusing ourselves or other scholars.
 - (iii) At some later stage of our study, we can revisit the issue and see if there is indeed any evidence for the 'VP analysis' over another analyses. And, if another analysis ends up being better, we can/must adopt that one.

"Any linguistic description is in reality going to exhibit certain kinds of theoretical biases, many of which may be implicit." (Crowley, *Field Linguistics: A Beginners Guide*; p. 17)

"A careful fieldworker... will also be constantly questioning the reliability of his or her own representation of that data, as well as his or her interpretation of what has been recorded." (Crowley, *Field Linguistics: A Beginners Guide*; p. 134)

(25) **Don't Be Afraid to Be Wrong, Even About *Phonology*!**

The advice above especially extends to the realm of the phonological. Sometimes, languages make use of phonological contrasts that take the average English-speaker *months* to detect reliably.

Thus, in a class of this sort, don't worry about getting the transcription 'wrong'. Always go with your 'gut instinct' regarding what you're hearing on the recording!

a. On Using Dictionaries / Grammars to 'Correct' Your Transcription

Sometimes, if you have access to a dictionary or grammar, there will be the temptation to use those resources to 'double-check' or 'correct' your transcription. *My advice is to avoid that temptation!*

- (i) Sometimes, earlier dictionaries & grammars can be mistaken.
- (ii) Sometimes, earlier dictionaries & grammars are actually describing a dialectal variety different from the one spoken by the consultant.

Your transcription of the recordings should be a record of what you, the linguist, hear on the recording.

"There is probably no linguistic researcher alive, nor will there ever be, who can write down at first hearing everything which is linguistically significant..." (Crowley, *Field Linguistics: A Beginners Guide*; p. 131)

"Even if you have managed to work out the phoneme inventory, it is often possible to mistranscribe individual words containing sounds that you would never have imagined you would find any difficulty with." (Crowley, *Field Linguistics: A Beginners Guide*; p. 135)