

7. Transcribe, parse, work out what you don't know, discard hypotheses which are no longer tenable and make up new ones.
8. Expect to do this many times over the course of a field trip.

To make sure that you have examples of many syntactic phenomena, you can go through a checklist, such as the one in Appendix D, or make your own list of things that you want to investigate. You could start with the table of contents of a grammar of a related language. Make this list before you leave for the field; you will certainly add to it and refine it while you are working actively on the language. Even if you don't invent your questions until the night before you ask them, have a good idea of the topics you will investigate before you leave.

### 6.5.2 Further reading

- **Documentation/Description:** Ameka, Dench and Evans (2006), Gippert, Himmelmann and Mosel (2006), Shopen (1985).
- **Elicitation:** Kibrik (1977), Payne (1997), Samarin (1967).
- **Interviewing:** Patton (2002: ch. 7).
- **Stimulus materials:** Bouquiaux and Thomas (1992), Vaux and Cooper (1999).
- Further information is available on the web site.

## 7

# Further Morphology and Syntax

In this chapter, I address further topics which arise in working with morphological and syntactic field data.

## 7.1 Elicitation of paradigms

As your collection of elicited sentences grows, you will acquire a collection of inflected words. However, you may not know what the citation form for all of these words is, and you probably won't know all of the different ways that they inflect. There will be lots of gaps in your data. Therefore, fairly early in your fieldwork, it is a very good idea to bring some systematicity to this data collection.

The first thing you need to do is to work out what you don't know. The easiest way to do this is to compile a list of all of the inflected words in your data. That way, you will know what words you have with which morphology. This also serves as a handy check for variant transcriptions.<sup>1</sup> Next, work out which data are missing. For example, you could compile a table of the nouns in your data arranged by inflection. It might look something like this table of data from the Cushitic language Qafar (Hayward 2001:635).

	<i>nibd-</i> 'awake'	<i>duf-</i> 'push'
1sg	<i>nibd-iyoh</i>	
2sg		<i>duf-tah</i>
3sg masc	<i>nibd-ah</i>	<i>duf-ah</i>

3sg fem		<i>duf-tah</i>
1 pl	<i>nibd-inoh</i>	
2 pl	<i>nibd-itoonuh</i>	<i>duf-taanah</i>
3 pl		<i>duf-aanah</i>

Now, you need to go about filling in the missing cells in the table. You may not be able to ask directly. Questions such as 'what is the ablative of *house*?' or 'what is the third-person singular feminine imperfect of the stem *nibd-*?' are seldom productive ways of talking to non-linguists. There are two main ways to get this data. Either make up a sentence to translate the word in a context which is likely to produce the relevant form, or make a guess as to what you expect the word to be based on your previous data and see if it's right. For the example in the table above, you would want to ask something like this:

- (16) I am pushing.  
 We are pushing.  
 You are awake.  
 She is awake.  
 They are awake.

You may also want to have a go at eliciting full paradigms, without waiting for words to come up in context. The easiest way is to collect a set of words in a simple frame.

- (17) 'He Xed today.'  
 He worked today.  
 He smiled today.  
 He saw a cat today.  
 He went to school today.  
 etc.

Then, alter the frame slightly, for example

- (18) 'He will X tomorrow'

People often find this type of task repetitive and boring. Making up stories can make the elicitation a bit less tedious. For example, to collect examples of allative-case marked nouns you could make up a story where someone goes to a series of places. A child has lost a favourite toy and they go *to the shop, to their grandmother's house, to the bus stop, to the car, to school*, and so on in order to look for the toy. Stories where someone has lost something are great for locational phrases in general, for example the child could look *under a stone, in the rubbish bin, on top of the table, beside the bed, next to the tree*. Such elicitation forms also make good children's books.

When checking forms in paradigms in isolation, however, check the forms of the same word, together. Don't do all past tenses, then all present tenses. It seems to cause confusion. Don't try to work too fast either, and only ask for one form at a time. There's a strong temptation to ask for several forms at once, but it often results in confusion. Avoid framing questions like 'How do you say "he is singing" and "I am singing"?' Your consultant may catch on to what you are eliciting and will give you the paradigm without further prompting after a while.

Once you have a good idea of the way words inflect in the language, when you come across a new word, do spot-checking of paradigms to make sure of what morphological class it belongs to. You will need to become adept at parsing on the spot. Work out what forms you need to determine paradigm membership early on in your analysis (if possible); for example, for knowing the full conjugation of a Latin verb, you need the infinitive, the past tense, and the passive participle; all other forms can be predicted from those three. Such key forms are called 'principal parts'.

## 7.2 Productivity

Not all morphology is productive, however, data on the extent of productivity will take a while to gather. You'll need some familiarity with the language to know what you need to test for. Productivity is partly associated with frequency. That is, if you come across a morpheme frequently, on many different items in the same word class, that's a good indication that it's productive. However, some morphemes might be rather rare, but still productive, and there may be some morphemes which occur on high-frequency items but are not generally productive.

To test productivity, come up with a sample set of words which you think might be able to take the morpheme and see if the resulting forms are grammatical. If they all are, it's not completely safe to assume that the morpheme is productive. You might just have been lucky in the words that you chose. If only some of the items are grammatical, see if you can see what they have in common. Is it a semantic class, or is it something to do with the phonological shape of the word, or is it something else? You can also ask your consultant to think of other words that have that ending on it. You might need to give some examples, rather than asking about the morpheme directly. Wug-testing also provides clues to productivity. Make up some words and see if consultants will accept derived or inflected forms (don't forget to check what they would mean).

It's very tempting to ask about productivity directly. However, questions like 'can this ending go on all words, or just on some?' can be very misleading. For example, your average English speaker when confronted with this question in the street ('are there lots of words in English with the ending -ness, such as kindness and happiness?') may either answer 'yes', depending on whether they happen to be thinking about Anglo-Saxon adjectives, or 'no', if they happen not to think of a couple of words off hand. People seldom have reliable explicit knowledge about productivity or frequency.

### 7.3 Selected topics in morphology

#### 7.3.1 Derivational morphology

Not all languages have extensive derivational morphology, and some have none at all. Testing for topics such as number marking is quite easy if you're using a contact language which has this type of morphology. It can be tempting to add a numeral or quantifier to make explicit that you want a plural item, but this may distort your data. Some languages only mark plurality when there is no numeral or quantifier present in the phrase. The language may mark more distinctions than the contact language. For example, some languages have both a dual and a plural, or a paucal. In some languages, paired body parts are always dual, while in others they are exceptions and take plural marking. Be on the lookout for a distinction between mass nouns and count nouns too.

You should also test for definiteness marking. It's best to do this in context, and to ask questions of your consultant about the implications of a given sentence. Do not expect the translated sentence to necessarily have the same definiteness values that you implied in the prompt.

It can be quite difficult to work out markings for definite and indefinite noun phrases. It helps to elicit them in the context of other quantifiers. It also helps to start with deictic marking, using the items that you can see and manipulate.

Eliciting derivational morphology which results in a change of word class can also be done through the contact language. It's easiest to do this if you point out the relationship between roots explicitly, rather than asking for words in isolation. You might say 'English has all these words which describe people doing actions: a *runner* is someone who runs, a *singer* is someone who sings, and so on. Can you think of similar words? For example, I remember that *sing* is *dar'taryun*. What's the name for a singer?' Sometimes this technique results in paraphrases, so you might need to ask specifically if there is a single word. A list of common derivational categories it is included in Appendix D.

#### 7.3.2 Class, gender or category marking, and conjugation membership

Some languages have gender systems, or inflectional classes for open word classes. There are strong a-real tendencies, so if you are studying a language in an area where noun class systems are common, you should be on the lookout for them and test for them. Noun classes often arise with numerals, with inflection marking, and on accompanying adjectives or modifiers. Therefore, you can test for them by using the same modifier with nouns of different shapes, size, texture, animacy or other feature. For more information about such systems, see Corbett (1991) and Aikhenvald (2000).

The discovery procedure for inflectional classes is exactly the same as for other morphology. That is, get examples of different words in similar syntactic contexts, and see how they differ. Keeping track of different inflectional classes can be quite difficult if the system is complex. The easiest way to do it is to compile charts with all the morphology, and keep a list handy of which words appear to belong to which class. Memorize as much as possible. Number the classes and include that information in your lexicon.

#### 7.3.3 Suppletive morphology and accidental gaps

Languages often have suppletive morphology (such as English *go* versus *went*). Suppletive pairs are easily missed early on. You will have so many gaps in paradigms that it might be very difficult to tell whether potential suppletion is that, or a word with similar meaning in a paradigm that you haven't yet fully recorded. The problem of identifying suppletion is

compounded by the presence of accidental gaps in paradigms. That is, a particular word might not have an inflected form in a particular cell of the paradigm. Over time you may notice systematic gaps in certain paradigms. For example, in Khoekhoegowab (also called Nama), paired body parts take singular and plural morphology, but no dual marking.

The studying of paradigmatic gaps is not something to be concerned with early on in elicitation. It is difficult to ask directly about paradigmatic gaps, since it is easy to accidentally 'coerce' a speaker into producing forms through overgeneralization. The best way to test for gaps is to get speakers to do a survey-like task where they are required to produce suspected gaps in a sentence. If there is a lot of disagreement in which form is produced, that is a good indication that there is a gap.

If this is common in the language you are studying, you should develop a way of marking genuine gaps from data gaps in your corpus.

#### 7.3.4 Case marking

Basic information about morphological case is easy to elicit. Ask your consultants to translate a set of sentences with different argument structure arrays (varying the semantic roles and grammatical relations). Take note of case frames, including any irregular case frames for certain lexical items (e.g., place names frequently have zero locatives in some parts of the world, or if your language has ergative case marking, there is likely to be a split somewhere in the morphology. In that case, it is quite easy to elicit sentences which do not give you meaningful information about the case frames of verb). For example, if the pronouns in the language are marked on a nominative/accusative pattern, and nouns on an ergative/absolutive pattern, subject pronouns will not tell you if the verb is transitive or intransitive if no object is present.

Going beyond the basics, however, is trickier. Your sample sentences are just the beginning of more advanced exploration of case semantics. You should test for optionality of marking, and possible interaction with constituent order. For example, in a good number of languages, overt marking of direct objects is optional, but the presence or absence of the case marker has effects on where the argument can appear in the clause, and whether it is obligatorily interpreted as definite.

#### 7.3.5 Tense and aspect marking

There is more to tense marking than elicitation of basic paradigms.

main clause. Present tense may be used in narrative to give an immediacy to the narrative, and tense, aspect and Aktionsart (roughly, the temporal structure of the event denoted by the verb) may interact in different ways. For example, in Yan-nhañu, the 'past' tense forms are also used to mark resultant states which may persist to the present (cf. *mukthun* (present) 'to be quiet, to become quiet'; *mukthana* 'to have become quiet, therefore to be quiet'). Be aware of the relationship between aspect and Aktionsart and design your elicitation questions to include verbs of different event structures. If you see an unusual example of TAM marking in a text, tag it and see if you can find similar examples.

Be aware of the role that aspect can play in elicitation. For example don't try to do aspectual elicitation on a verb such as *die*, which is often irregular.

Tense can interact with temporal marking elsewhere in the clause, so be sure to test verbal tense marking both with and without temporal adverbs or other temporal marking.

#### 7.3.6 Variation and optionality

It is not uncommon to find considerable apparent variation in morphology. Perhaps you have examples of a suffix which appears to alternate in free variation with another suffix. For example, you may have an antipassive construction where the demoted object can be either in the dative case or the locative. Or, you might have examples where a particular morpheme appears to be optional, and its presence or absence does not appear to affect the meaning of the clause.

Such variation may be due to several factors: perhaps different speakers have different preferences where more than one choice is possible. Different speakers could have understood your prompt question in different ways. Alternatively, the difference in morphological marking could be due to age-based or other dialectal variation. There might be genuine variants which are conditioned by a factor that is not obvious from the data you have so far. Finally, it could be an error of some kind, either a speech error, a transcription error or something else. Usually, the only way to sort out such variation is to wait until you have more data.

Make a note of variation in your data early on. It's worth coming up with a tagging system for your notes or a set of keywords that you can

### 7.3.7 Discourse-based morphology

Some types of morphology are almost impossible to elicit in a planned fashion. You will come across them as language is used spontaneously. This is yet another reason why it's a very good idea to learn to speak your field language, and to listen for the way language is used around you even outside of structured elicitation sessions. One example of this is the way that morphology and intonation interact in vocative constructions. In many Northern Australian Kimberley languages, when someone is calling out to someone they lengthen the vowel of the last syllable of the person's name (or the word), and they add a suffix *-w*, which may itself be held for a considerable length of time. In transcriptions, this might be represented as *kaliyawwww* 'Byeeee!' (in Nyikina). directly eliciting something like this would be almost impossible, so make sure that you are aware that you cannot conclude from elicited data what categories are not present in the language. Other similar discourse-based morphology will, of course, only turn up in discourse.

### 7.4 Handling unknown morphology

It's not uncommon in the first few months (or years!) of fieldwork to be constantly assailed by new morphology, or by new uses of familiar morphemes. In early elicitation, it is also very common to miss morphemes completely. That is, you will treat as morphologically simple words which are, in fact, morphologically complex. It is also possible to over-generalize, and to find apparent morpheme boundaries in words that are actually simple. There is no way to solve this other than to increase your knowledge of the language.

If you don't know what a morpheme does, ask your consultant for other sentences which contain the inflected word (e.g., if I have an example of *nhälil*, and I know what *nhä* means, but not *-lil*, I might ask 'Could you tell me some other sentences/things to say which have *nhälil* in them?'). Try to generalize the morpheme to other words and ask your consultants to put them in sentences. If you are not sure if the word is morphologically simple or morphologically complex, try to inflect the word yourself and ask your consultants to put it in a sentence. If the word is actually simple, your form will probably sound to them like a language joke. For example, for some speakers of English the

wanted to test whether integral did in fact have this prefix, you might ask whether there was a word *intragral*.

In other cases, you might have identified a morpheme but you might not be sure what conditions the allomorphy. Your data might be misleading. Here is an example from my field-methods class in 2005–06. Eastern Armenian has two nominal plural allomorphs, *-er* and *-ner*. Based on our class data it was possible to say (with a few exceptions) that *-er* was used for consonant-final stems and *-ner* was used for vowel-final stems. On the other hand, the data also supported an interpretation where *-ner* occurred on words of more than one syllable, and *-er* on monosyllables. In fact, the exceptions were covered by this latter rule. Subsequent testing showed that the number of syllables is the important conditioning factor. This illustrates the need to revisit hypotheses if you start finding exceptions. While many rules have exceptions, does not mean there is not a better generalization.

### 7.5 Common problems

There are some problems that come up exceedingly commonly in analysis of previously undescribed languages. One is clitics. Is a given form a clitic or an affix? You might have some evidence that points one way, and some that would lead you to the opposite conclusion. The important thing to do here is to work out what evidence you have one way or another, and then try to work out what it would take to decide in favour of a particular analysis. What would you need to know in order to discover which of your two hypotheses is correct?

It is very common to misparse words. It is very easy to be misled along particular morphological analyses early on in the description of the language. Expect this to happen. It is the result of not having enough data (or in some cases, either having too much data, or not having enough of the right type of data). Don't feel discouraged, the most important thing is to be able to keep track of your analysis in your head, and to recognize when you get information that confirms or contradicts it.

It is very tempting to regularize transcription to your citation forms. You may do this, however, at the risk of overlooking conditioning and regularizing your data. Be very careful. This is why it is a good idea to make notes on realization if you're transcribing phonemically.

It is also very easy to set up a pattern and then to have consultants

pattern. Try this yourself by asking an English speaker for the past tense of the following verbs:

- (19) sing      (sang)  
       ring      (rang)  
       ping      (pang)

You can avoid this problem by not doing too much elicitation at one time, by using the forms yourself in sentences, by eliciting with two people there to check each other, or by randomizing forms in paradigms. (I don't recommend the last one, as it is confusing and tiring to consultants.)

The biggest problem that tends to happen with elicitation is that it is difficult to realize the degree of reliance placed on existing structures in the analysis. That is, it is quite difficult to move away from assumptions from the languages you know well to see what it is that is going on in the language you are studying. There is a strong default assumption that things will be the same unless you have evidence to the contrary. Therefore it is very easy to overlook potential differences, simply because you assume they will be the same as your native language.

## 7.6 Further topics in syntax and semantics

### 7.6.1 Constituent order

Early on in your elicitation you will get data on constituent order. However, it is quite easy to be misled by elicitation on this topic. For example, if your language has no basic word order (i.e., discourse determined word order as is found in many languages in Australia and native North America), the same order will probably appear in the prompt and translated sentences. Therefore, if you're using English as your metalanguage and all your responses are SVO, you cannot be sure that SVO is the basic word order of the language. You should try early on to move constituents around in the sentence. Another common pattern in free word order languages is for common discourse participants to receive topic position in sentences. That is, the elicitation itself creates a discourse context.

Even if you do not work in the syntactic theory that generates sentence structure from rules, it is a good exercise to try to describe your data in those terms. It gives you a good idea about potential gaps in your data, possible ambiguities, and allows you to make predictions about what possible and impossible sentences might look like.

Some progress can be made in constituent order discovery through elicitation. That is, you can take a sentence, switch the order of constituents, and ask if the result is still grammatical (and what it means). However, your best clues to constituent order will come from textual data.

### 7.6.2 Scope

**Scope** is a core topic in the generative literature. However, it can be very difficult to elicit scope judgements, or even to get people to see that there's a difference in scope readings. For example, a non-native speaker of English might not realize without prompting that 'every child caught a fish' is ambiguous between wide and narrow scope readings (one in which every child caught one certain fish, versus the reading in which each child of a set ended up with a fish, but not necessarily the same one). The best way to elicit these types of judgements is to set up pictures (or scenarios) and ask for translations. Do it in a small group and allow for the participants to discuss the possible meanings. Try to have something concrete to illustrate the different meanings if possible. For example, you could have a set of coins and ask about the sentence 'every child got a coin'. Have consultants illustrate whether such a sentence in their language implies that each child got a different coin, or whether they only received one, or whether either interpretation is possible.

### 7.6.3 Causative constructions

Not every language has causative constructions; some have both simple and complex causatives, and others have only one. Some languages distinguish direct and indirect causation. Causative marking might be morphological as a derivational device only, or it may be more productive. Some languages distinguish reasons from causes. That is, the reason for an action might be marked differently from the cause (or causer) of an action. Eliciting such constructions can be quite difficult at first. If the language has both an analytic causative and a morphological causative, eliciting through the English analytical (make) causative often primes consultant to produce analytic causatives. It is also quite difficult to get a sense of the semantics of such constructions through elicitation. If you're having trouble finding such constructions in the language, it may be worth to wait until you have textual materials. Or, you could use a set of stimulus materials which have causers in them, such as videos where one participant makes someone else do something.



### 7.6.4 Commonly missed constructions

Some constructions cause particular problems in field-methods classes, particularly constructions which are not very common in European languages and which are not covered in standard classes.

Numerals and number marking often interact. That is, it may be that number is only marked on a noun phrase when the numeral is present. This can cause intriguing conclusions, because a tempting way to try to elicit plural marking is through including a number. The numeral system itself is worth documenting, by collecting not only lower numbers and decimals but also larger numbers, and ways of describing multiplication, addition, subtraction and other mathematical operations.

Reflexive and reciprocal marking may occur even when there is no overt reflexivity of the action (see Gerdtz and Hukari (2006) for an example). Applicative marking and antipassives can also cause problems. Kroeger (2004) has a good description of applicatives. Copular clauses are often not investigated extensively. Remember to test for different types of predicates (e.g., with a nominal predicate, as in *my friend is a teacher*, as well as adjectival and other types of predicates). Tense marking and negation can also interact with copular marking.

It is very difficult to get translations of deictic markers, since they can be used not only spatially and literally, but also metaphorically in many languages. There is a tendency to assume that there will be a precise translation for *this* and *that*. It is very tempting to assume that the lexical category (or part of speech) of an item will be the same as it is in English. This is a very dangerous assumption to make. You will need to come up with your own tests for parts of speech in the language. Subclasses of nouns may behave in different ways with respect to case marking. Personal names in place names many inflect differently from common nouns (or may have a reduced number of case possibilities).

It is common to spend much less time on the syntax of languages that have complex morphology than on languages with simple morphology, and that can lead to a lopsided description. For example, some grammars of northern Australian languages have a hundred pages devoted to morphology, but only ten or so to syntax. Don't fall into this trap.

Another area that frequently causes difficulties is complex predication, or clausehood more generally. It may be initially very difficult to determine whether a given string of words comprises a single clause, or more than one. There are syntactic tests for this, however. For example, negation marking is clause bound; therefore if you have a single negation marker which takes scope over both verbs, that is a good indication that the verbs belongs to the same clause (for more tests, see Bower 2006).

## 7.7 Where to from here?

### 7.7.1 The next step

If you work this way with elicitation material, you will have a fairly good idea about some aspects of the language. You'll know how to form basic sentences, and quite a lot about less basic structures as well. You will have acquired a knowledge of various (more and less useful) vocabulary items. You'll have a lot of data which will be useful for phonetics, phonology, morphology and syntactic work. You'll probably also have a lot of unresolved questions. It's also important to realize what you don't have. You don't have much data about pragmatic factors of language use. You probably have a lot of items glossed 'particle', 'emphasis', or '??'.

The next stage is to expand your data collection methods. As discussed in the introduction to Chapter 9, elicitation doesn't give you the full story. Textual data will help to clarify some of these questions and will probably give you yet more puzzles.

The best way to explore topics for morphology is to have a good grasp of the literature on morphology, morphosyntax and typology. There are numerous overview volumes dealing with different morphological and syntactic phenomena, including Aikhenvald, Dixon and Onishi (2001), Comrie (1989), Corbett (2000, 2006). General morphology textbooks, such as Haspelmath (2002) and Bauer (2003) may also be useful here. Collect examples of the construction<sup>2</sup> and work out what you think the essential features of the construction are. Then go hunting in the literature for something similar (and ask colleagues and your advisor), then work out in what ways the construction you've identified is similar, and how it differs.

### 7.7.2 What if none of this works

It may happen that none of the methods described here are useful to you. Your consultants might not be comfortable working with elicited sentences, and may refuse to translate. You may not find anyone who speaks the language fluently, and so the materials you've prepared might not work.

In that case, perhaps your best option is to be guided by your consultants and learn what they can teach you. For example, if your consultants are not comfortable working with elicitation, they may have their own ideas about how to teach a language. That may still give you a great deal of data if you are patient and can extract the relevant generalizations from the materials you're given to work with.

Moreover, while I have stressed the importance of negative data, elicitation and grammaticality judgements, it is possible to do a great deal of work and get a great deal of data from positive materials alone.

### 7.7.3 Working in the field language

During your field trip it may become desirable or necessary for you to work in your field language. I have not included information on monolingual elicitation here – see Everett (2001) for more information. You may wish to work with monolingual as well as bilingual consultants, though. It's possible to do this even without fluency in language, and I recommend it highly as a way of helping you to increase your fluency quickly.

Your early aim in working this way should be simply to get people talking. You'll be able to go through what they said and transcribe with bilingual consultants later on. If you have some ability to respond in the language, you will probably get fairly simple structures, as people will use foreigner talk to you.

Some techniques you can use include asking very general open-ended questions as a way of encouraging discussion. You can also present stimulus materials. Asking for vernacular definitions is also very easy to do with monolingual consultants.

### 7.7.4 Further reading

- **How to set out an argument:** Harris (2000) is an excellent example; see also Green and Morgan (1996).
- **Auxiliaries:** Anderson (2005).
- **Semantics:** Evans and Sasse (2005), Matthewson (2004, 2005).
- **Quantifiers:** Enç (1991).
- **Experimental design:** Crain and Thornton (1998: chs 16, 17, 24).
- **Various morphosyntactic categories:** Aikhenvald (2004), Chafe and Nichols (1986), Comrie (1976, 1985), Dixon and Aikhenvald (2000), Dixon (1994), Lyons (1999), Palmer (1986).
- **Syntax:** Bouquiaux and Thomas (1992), Levinson and Wilkins (2006), Payne (1997), Shopen (1985), Thomas (1975).
- **Word order and (non)configurationality:** Austin and Bresnan (1996), Jelinek (1984), Kiss (1998), Mithun (1992, 2003).
- **Parts of speech:** Schachter (1985).
- **Monolingual fieldwork:** Everett (2001).
- **Typology:** Comrie (1989), Haspelmath, Dryer et al. (2005), Nichols

## 8

# Lexical and Semantic Data

In the previous chapters we have discussed sentences, morphology and grammar. However, we should also talk about the documentation of the lexicon of a language. Lexical documentation can be something done in conjunction with other work on the language, however dictionary making is also an extensive enterprise in its own right. You could just wait for lexical items to appear in the course of your other work (e.g., in text collection); however, the returns on this method diminish rapidly as you gain familiarity with the language (i.e., the number of new words in texts drop off rapidly).

## 8.1 Getting vocabulary

One way to collect words is just to ask for them. We discussed basic lexical elicitation in §3.1.3. Those methods can be extended to more detailed and less common vocabulary within various semantic fields. Collecting antonyms, converses and hyponyms at the same time as a lexical item can be useful. For example, when you ask for 'little', ask for its opposite at the same time. (Don't just ask for the English opposite, in this case 'big', ask for the 'opposite'.<sup>1</sup>) You may also want to branch out early and ask for 'very little', 'littler' and 'littlest'.

I've always found it easier to do vocabulary elicitation in small groups (three or four consultants). Consultants prompt each other and the arguments about definitions are usually interesting and good sources of conversational data. In a field-methods class this probably will not be feasible. Don't worry if your consultant can't remember a word, you can always come back to it later. It's easy to forget the word for something if not on the spot