1 Introduction
What is linguistic fieldwork?

Are you ready to 'go to the field'? This book is written to prepare you for a great adventure: to discover a language, to use the skills you have learnt studying linguistics to document, describe and analyse how people use language. Your adventure may take you far away to a place you never heard of before, or you may simply talk to your next door neighbour. Linguistic fieldwork can be many things!

Have you ever wondered why your neighbour, who speaks both Urdu and English, sometimes switches between the two within the same sentence? Or have you always wanted to work on an 'exotic' language in a far-flung place? Or are you interested in a particular linguistic theory and want to figure out whether a theoretical notion can be found in language? In all three cases, doing linguistic fieldwork would be a way of finding out. You could record people's speech, carry out linguistic experiments or ask the speakers questions about their language use. These methods can be used on their own or in combination with one another.

Apart from being one of the most satisfying ways to engage with linguistics, fieldwork is essential in expanding our knowledge of how language works. We would not be able to get new ideas by theorizing alone. Real data is needed to make valuable discoveries. In this way fieldwork is indispensable for theoretical development.

Fieldwork can confirm what you suspected or lead to surprising results. You may end up with ideas completely different from those you started out with. Answering one question during fieldwork may lead to a hundred further questions. Yet, it is enormously satisfying to eventually find out how the language works.

Finally, fieldwork can be time out from the other things you may do. While the idea of being an 'armchair linguist' (Crowley 2007: 11) may be appealing (it sounds comfortable!), having a change of scene can be a very good experience.

Linguists choose to do fieldwork for various reasons. The first reason may be a university assignment. Sometimes the main reasons are not linguistic. For example, one of the authors, Dan Everett, was originally a missionary. Fieldwork was his way of learning the language to be able eventually to translate The Bible. As it turns out, fieldwork has not only changed his views on Christianity, but has also altered his preconceived ideas about the grammatical theories he was following at the time (Everett 2008).

1

Jeanette Sakel

For me it all started when I was a student writing my BA dissertation. Living in Denmark at the time, I decided to carry out sociolinguistic fieldwork on the maintenance of the Greenlandic language among Greenlanders living in Denmark. It was primarily a questionnaire-based study, which some may argue is not 'fieldwork' at all. However, for me it was the beginning of my career doing field research. I discovered that I loved working with speakers on their language. Subsequently, I decided to collect data on aspects of the grammar of Greenlandic for my MA thesis. I successfully applied for a small university travel grant to go to Greenland, which paid for my airfare. My Greenlandic teacher in Denmark put me in contact with a *højskole*, a type of residential school for adults in Sisimiut, just north of the Arctic Circle on the west coast of Greenland. Paying a small fee, I was allowed to stay at the school. The setting was ideal for my study, as I had easy access to speakers of various ages and dialectal backgrounds. On the professional level, the fieldwork went very well: I collected interesting data and learnt more about the structures I was studying. But there was a surprise in store that I could never have anticipated, the details of which I will go into in section 5.2.9.

1.1   A definition of fieldwork

You will probably have an understanding of what fieldwork is and an idea as to the kinds of activities that take place. Before we go into detail defining our subject matter, have a short think about your understanding of the term.

Exercise 1.1

(a)   What do you understand by the notion of *fieldwork*?
(b)   How does your understanding relate to the following points:
–   interviewing somebody?
–   interviewing yourself?
–   working in a setting far away?
–   working with your next door neighbour?
–   working on an undescribed or little-described language?
–   working on a well-known language?
–   documenting a language or writing a grammar?
–   looking at a particular aspect of a language?

If we were to ask ten people how they define fieldwork, we would probably get ten different answers. Indeed, many linguists have diverging views, centring on a common core of 'prototypical' fieldwork, which involves collecting data on an endangered language in a remote and usually 'exotic' setting. Hyman (2001), discussing the ways in which the term *fieldwork* is used in the literature, distinguishes

prototypical and non-prototypical fieldwork at various levels. Prototypical field-work entails the linguist working with speakers in a small setting far away for a long period of time. The language is spoken in its natural language context, the data are naturalistic and the motivation for conducting fieldwork is entirely language-driven. The opposite of each of these points would be non-prototypical fieldwork, namely the linguists interviewing themselves or being observed by others in a large setting close to home. The fieldwork would only last a short time, the language under scrutiny being well known. The subject matter would be the formal system of a language, the language data controlled and the entire study driven by theory.

Indeed, many instances identified as linguistic fieldwork are not prototypical, but lie somewhere in between the two extremes.

Exercise 1.2

(a)   To what degree were the two types of fieldwork carried out by J. Sakel prototypical or non-prototypical?

(b)   Imagine the different fields of linguistics that use language data – what type of fieldwork would they require to collect their data?
Discuss, for example, the type of data-collection needed in fields such as generative grammar, sociolinguistics, typology, documentation of endangered languages, second language acquisition and bilingualism.

Depending on your background in linguistics, you may agree with the prototypical definition of fieldwork. However, many linguists refer to their work as fieldwork even though it is far from prototypical. Where should we draw the line between what is fieldwork and what is not? Imagine a generative lin-guist, who often uses introspection to theorize about grammatical structures in English. Would this person be conducting fieldwork by doing so? How many people would say that the native-speaker grammaticality judgements of the early generative grammarians (e.g. Chomsky 1965) are indeed fieldwork? As a matter of fact, this type of work almost totally fits Hyman's (2001) categorization of non-prototypical fieldwork.

Take, for example a variationist sociolinguist like William Labov, who works on varieties of English in the USA (e.g. Labov 1966). He is not only close to home, but also working on a large, well-described language which is also his native tongue. Still, often linguists refer to his work, as well as others' work in sociolinguistics, as fieldwork (see also Milroy 1987 for the use of naturalistic data in sociolinguistics, which again is usually referred to as fieldwork).

Consider then the typologist Marianne Mithun, who conducts fieldwork on a variety of endangered North American indigenous languages, many of which are spoken close to her home in California. Even though most people would agree that this should be considered fieldwork, it does not fulfil all of Hyman's (2001)

prototypicality criteria, for example she may indeed work with her neighbours on a particular structure of a language.

Finally, let us consider the fields of second language acquisition and bilingualism, which rely heavily on the analysis of actual language data, but where research sometimes involves experiments in controlled settings. Furthermore, at least one of the languages is usually well known. Language data in these fields can range from natural to artificial, spontaneous or semi-spontaneous speech, and researchers may be using elicitation or conducting experiments (Nortier 2008: 46). Still, many researchers in this area refer to their data-collection as fieldwork, in particular if it involves primarily naturalistic data (see Moyer 2008: 27 and Dörnyei 2007: 130).[1]

As you can see, very different disciplines within linguistics conduct fieldwork, some more 'prototypical' than others. For this reason, definitions of what fieldwork *is* vary considerably. In order to reach our definition of fieldwork, let us briefly acknowledge the history of linguistic fieldwork and how others have defined fieldwork before us.

Linguistic fieldwork as a discipline began with Franz Boas (1858–1942), who trained a core of linguistically aware anthropologists and thereby was (indirectly) responsible for the birth and growth of North American linguistics. During the years of Boas's influence, roughly during his life and following his death until the 1950s, North American linguistics was concerned with describing specific languages in detail, producing integrated studies of texts keyed to cultural studies, grammars and dictionaries. Other influential forerunners were Boas's student Edward Sapir, as well as Leonard Bloomfield. For a more detailed discussion of the history of fieldwork in the Americas, see appendix 1.

Over the following decades, a number of publications appeared to help the linguist or anthropologist learn and figure out the grammar of a language. These were guides on how to extract information about grammar from spoken language data (e.g. Nida 1947), sometimes focusing on particular language groups (e.g. Bouquiaux & Thomas 1976; see also later works on particular language groups, e.g. Abbi 2001). A notable exception was Samarin's (1967) book on field methods, which also gave information on more practical issues, such as what makes a good language teacher. Samarin's view of fieldwork is rather broad, e.g. regarding the place of study: 'Field linguistics can be carried on anywhere, not just *in the field*, as its name implies. [emphasis Samarin's] A "field archaeologist" must go out to where he expects to collect his data, but a linguist can bring his data to himself. Thus, some fieldwork is done by bringing jungle dwellers to a city and is conducted in an office instead of a lean-to.' (Samarin 1967: 1–2).

As interest in linguistic fieldwork has increased, a range of publications on modern fieldwork methods have appeared (e.g. Vaux & Cooper 1999; Newman & Ratliff 2001; Crowley 2007; Gippert, Himmelmann & Mosel 2006; Aikhenvald 2007; Bowern 2008; Chelliah & de Reuse 2011). There is also a plethora of online resources on fieldwork. These published materials present fieldwork methods, including modern concerns such as fieldwork ethics. Still, the type of fieldwork focused on is in many cases close to the prototypical 'ideal'.

Newman & Ratliff (2001: 1) refer to fieldwork as 'the complex and involved business of describing language as it is used by actual speakers in natural settings'. The focus on 'natural settings' means that working in an office, as mentioned by Samarin (1967), would not count unless that is where the speakers generally use their language.

Crowley (2007) sets fieldwork ethics high on his agenda. He focuses on the documentation and description of endangered languages and views fieldwork as a means to record the languages that are in danger of disappearing. Crowley acknowledges less prototypical work, e.g. with only one speaker as part of a university field methods course, as a '*kind* of fieldwork' (Crowley 2007: 14, his emphasis), agreeing that 'Any work that you do with a speaker of a language other than English with a view to publishing the resulting linguistic analysis can legitimately be referred to as fieldwork.' (2007: 14). However, he is reluctant to endorse this type of work. Similarly, while mentioning sociolinguistic data-collection, he does not include this in his definition of linguistic fieldwork (2007: 18).

In a similar way, Bowern (2008: 7) stresses the humane and ethical perspective of fieldwork, which also shines through in her definition of fieldwork: 'It involves the collection of accurate data in an ethical manner. It involves producing a result which both the community and the linguist approve of. [. . .] The third component involves the linguist interacting with a community of speakers at some level.' According to Bowern, linguistic experiments would not be considered fieldwork *per se*, neither would working with a speaker removed from his language community.

Chelliah and de Reuse (2011) focus on descriptive fieldwork and language documentation, but are more inclusive than other publications in asserting that fieldwork can be carried out in one's home town, among one's own relatives. This is to cover cases where native-speaker linguists work on their own language. Nonetheless, the general focus of the book is on prototypical fieldwork.

Our own definition aims to encompass both prototypical and less prototypical types of fieldwork:

> **Fieldwork** *describes the activity of a researcher systematically analysing parts of a language, usually other than one's native language and usually within a community of speakers of that language.*

It is useful to consider this definition in more detail. Again, by dwelling on our own definition, we are not claiming that it is 'right' in some absolute sense. But it does raise issues worth considering, however one ultimately comes to understand the essence of field research.

'Systematically analysing' should be clear. We go into field research with a system of ideas that guide our research. How does this system guide, then, what we are going to study? What are the subparts of the system? How do the different subparts of the system, projected onto the language of study, interact? For example, perhaps we are conducting research to test a specific claim in the literature, e.g. 'language **x** lacks embedding'. What system could there be to

our investigation? First, if a claim has been made to this effect, we want to check with native speakers the data that were adduced on its behalf. Do native speakers agree with all the grammaticality judgements offered to support the claims being tested? Are there discrepancies across speakers? And so on. To check data requires a plan. How many speakers should one check the data with? How should one subcategorize and test discrepancies in speaker judgements? How can one design and test alternative hypotheses? Second, if a language lacks embedding, it should be reflected at several places in the grammar, not merely, say, in the absence of complement clauses. Does the language have disjunction? Coordination? Verb phrases? Complex noun phrases? Adjectival phrases? And so forth. In other words, field research is like any other large, complex task. It requires planning, administration, progress checks, self-evaluation and reports (at least to oneself).

Continuing with a discussion of our definition, why does it refer to 'parts of a language', rather than a 'whole language'? First of all, it is impossible to study a whole language. Just consider the thousands of studies of the English language and the fact that there is no sign that research on English is coming to an end. A language is vast and beyond any single researcher's ability to study in a human lifetime. Language is everything: semantics, sociolinguistics, phonetics, phonology, syntax, morphology, ethnography of communication, and so on. Second, language as an object of study is unclear, unfocused – there are no boundaries to identify either a coherent beginning or end of the study if its object is 'documentation of English' or some such. Additionally, languages change, and in that way no language can be studied conclusively. The goal of 'parts of language' requires a lowering of the sights from *language* to *selected components*. Their selection requires a coherent vision of how the parts fit together, assuming that the study is to fit together at the end, that it is not strategically opportunistic.[2] Having said this, linguists are usually aware of this fact, and when they say that they are 'writing a grammar of a language' they often do so because it is easier to say than 'writing a grammar of parts of the language, the way we understand the language at this particular time, with the structures found in the corpus that we managed to investigate, considering that we may not have been able to record every last speaker's idiolect of the language, etc.'. The linguist could resort to saying that they are 'writing part of a grammar of a language', but that may be understood as focusing on certain aspects (e.g. the verbal paradigm), rather than doing a general overview. Talking about a 'sketch' of a grammar may also be an option, but such sketches can be of a few pages, while the linguist may aim to make the grammar as comprehensive as possible. For the sake of simplicity, in what follows we refer to 'writing a grammar' or 'documenting a language' with the understanding that one can only ever record 'part of a language'.

Coming back to our definition of fieldwork, the last two parts of the definition state that something would 'usually' be the case. By this we try to delimit field-work from other types of data-collection, while not excluding non-prototypical

forms of fieldwork. For example, the language under scrutiny would in many cases be other than one's native language. Yet, many modern fieldworkers train speakers to conduct fieldwork on their own language, or speakers study linguistics to eventually be able to work on their own language. The work carried out by these speakers would still be considered fieldwork. They would probably use the same methods as other fieldworkers, for example not relying too heavily on introspection but analysing the language of a variety of other speakers. Indeed, this type of fieldwork is very rewarding, as the speakers' insight into their language may lead to further insights into the language that a non-native linguist may not spot when studying the language 'from the outside'.

The last part of the definition states that field research would usually be conducted 'within a community of speakers of that language'. This means in the place and among the people where the language is usually spoken. Thus, you could be living among an indigenous group far afield, or if you study the language behaviour of your neighbour mixing Urdu and English, that community may merely involve the neighbour and his family.

Let us briefly go back to Samarin's (1967) statement that fieldwork can also be conducted away from the speech community. Indeed, the locus of language, the speaker, can easily move to a different place and therefore a linguist is not as restricted as an archaeologist when it comes to the place of work. Indeed, in many cases, circumstances make linguists work elsewhere than in the original speech community. Imagine for example a language spoken by refugees, whose original home is torn by war. As linguists, we can still study the language by bringing one or more speakers to us. But will we miss out on something? By working with only one speaker, we may not capture the entire language, but rather document an idiolect. Think, for example, about the way in which young speakers differ from old speakers in their language use. Furthermore, we may not be able to check the data with other speakers. If our speaker has contact phenomena or idiosyncratic behaviour in his language (such as a lisp), we may not pick up on it and may instead assume this is how the language is spoken by others as well. On the other hand, imagine we had access to a range of different speakers in our setting away from the original place where the language is spoken. This would open up the opportunity to be able to check the data, but there might still be issues that could have been different had the fieldwork been done *in the field*. If we assume that language and culture are inextricably intertwined (Everett 2005), one would not be able to understand one without the other. Taking speakers out of their communities or studying parts of languages outside their cultural contexts may lead to very different results from studying a language in its natural environment. The language may not be the same as it is when being used on a regular basis. Grammar and culture can affect, and to some degree effect, each other. Coming back to our definition: where fieldwork is possible within a community of speakers, it is a good idea to conduct as much fieldwork as possible in that community. Nonetheless, there are cases where it is necessary or advantageous for fieldwork to take place in other settings.

Summarizing, our definition is kept rather broad in order to include various linguistic pursuits dealing with empirical data. It encompasses prototypical fieldwork on the grammar of a language, as well as fieldwork in areas such as sociolinguistics and bilingualism.

1.2   Overview: the following chapters

The chapters of this book deal with the different considerations surrounding any type of fieldwork – hence every chapter will be relevant on its own account. While we include quite diverse types of fieldwork, some sections within individual chapters may be more relevant to you than others. As most subsections are generally readable on their own, you can easily pick those you are interested in, skipping less relevant aspects.

In chapter 2 we present you with two examples of fieldwork, with hands-on exercises to try out some of the techniques. You will first learn about the basic skills of text-collection, transcription, analysis and elicitation to study the language Mosetén from the Bolivian Amazon. Then, you will be introduced to how to set up a fieldwork project on language contact between Somali and English among members of the UK Somali immigrant community in Bristol.

Chapter 3 focuses on the languages involved in fieldwork. We discuss ways of finding a language to work on, lingua francas, as well as the ins and outs of monolingual and bilingual fieldwork.

Chapter 4 deals with the people involved in the fieldwork, including the researcher, the speakers of the language or variety under scrutiny, as well as other stakeholders. We discuss how to find appropriate speakers and we go into detail concerning fieldwork ethics.

Chapter 5 is about the preparation of the fieldwork. Our discussion ranges from preparing your research questions and literature review to applying for funding and preparing for fieldwork in remote places.

In chapter 6 we go into detail on a wide range of fieldwork methods, including the collection of texts, elicitation, linguistic experiments and participant observation.

Finally, chapter 7 examines the outcomes of fieldwork, including corpora, grammars and dictionaries. We also discuss how to archive fieldwork data.

1.3   Summary and further reading

In this chapter we introduced the concept of linguistic fieldwork and discussed our definition of the subject. For further reading on the distinction between prototypical and non-prototypical fieldwork, see Hyman (2001). For a

detailed overview of the history of linguistic fieldwork with a wealth of references to earlier works on linguistic fieldwork, see Chelliah & de Reuse (2011). The latter book is also a modern, general introduction to fieldwork, alongside Bowern (2008) and Crowley (2007). A book aimed at a general readership, which discusses numerous aspects of linguistic fieldwork among the Pirahã, is Everett's (2008) *Don't sleep, there are snakes*.

2 Fieldwork projects
Two examples

This chapter guides you through aspects of two very different fieldwork projects:

1. Fieldwork on the morphosyntax of an Amazonian language from Bolivia: how to begin extracting information about the grammar from spoken language

2. Fieldwork on the language contact situation between Somali and English among immigrants in the UK: how to set up a research project, and what to take into account before starting the data-collection

This is a first opportunity to learn about – and try out – some of the fieldwork techniques discussed in this book. It will be an introduction to many of the techniques and terms discussed in later chapters. These include text recording, elicitation and transcription in **1.** and how to find speakers, use picture stories, use sociolinguistic questionnaires, apply for funding and ethics applications in **2.**

2.1 Fieldwork project 1. Mosetén in Bolivia: text-collection, transcription, analysis and elicitation

In this chapter we will introduce the basic principles of text-collection, transcription, analysis and elicitation. These are the cornerstones of 'prototypical fieldwork', generally employed in describing the morphosyntax of a language in order to write a grammar, but aspects of these methods are present in all types of fieldwork.

The data are taken from Sakel's fieldwork on Mosetén in Bolivia, a language spoken by approximately 800 people in the Alto Beni region of the foothills of the Andes and the Amazon Basin. Sakel's aim was to write a grammar of this language. This is how she spent her first weeks in the field.

Jeanette Sakel

When I first started working on Mosetén I sussed out people who would be willing to work with me and who would be good teachers. Within my first week of living among the Mosetenes I was contacted by Juan Huasna, a Mosetén man in his late forties who had been working with anthropologists and biologists doing fieldwork in

---

10