

linguists. This raises a number of methodological issues including that of the status of ethnography in linguistic research. There is a possible unintended drawback to the requirement for research teams to be interdisciplinary. It may reproduce an anachronistic division of labour, allowing the representatives of the two disciplines to sit back and leave things as they are (or used to be), namely by considering ethnography to be something that concerns anthropologists only. Therefore, one of the first questions to deal with is as to why linguists should bother about ethnography given that there are so many other things to worry about. Secondly, what exactly do we mean by ethnographic material? And thirdly, where and how do we place ethnographic data in the archival record? The present article deals with these questions. It is based on a presentation given at the most recent DoBeS training workshop and is published here with the hope that reports on new ideas and experiences concerning ethnography in the course of joint work carried out in language documentation projects will become a regular feature of this newsletter.

Why bother about ethnography?

If language documentation is to serve as a basis for long-term archiving it needs to be “meta-theoretical” (the term used by the VW-DoBeS programme). “Meta-theoretical” is not to be confused with “non-theoretical” because strictly speaking every documentation has a theoretical frame, however implicit it may be. Rather, “meta-theoretical” translates into “amenable and useful to a broad spectrum of researchers and theoretical interests”. This goal cannot be achieved if data collection follows solely the current, and to some extent arbitrary, interests of the researchers involved. Moreover, the problem is not only one of *what* kind of data should be collected but also one of *how* it should be organized for the archive, in other words how it should be “chopped up” into sessions as searchable units of such an archive. The problem may be approached by orienting the research less according to the positions of researchers in their contemporary intellectual landscape and more according to the positions of speakers in *their* social landscape. The research strategy that promises to open up this social landscape to us is ethnography. In other words, ethnography helps to reduce the arbitrariness in data collections by drawing on the cultural context of speakers and it helps to make the language documentation materials meta-theoretical enough to be suitable for a long-term archive.

In researcher-directed data collection, as for instance in structured interviews or stimuli-generated contexts, data sessions are demarcated by decisions of the individual researcher - however arbitrary these may seem to others - as to what and who is to be included in the record. When moving away from this established, narrow mode of data collection the question of what constitutes a session has to be resolved through ethnography. Ethnography suggests an orientation

Ethnography in Language Documentation

Thomas Widlok

MPI, N megen, DOBES project Akhoe Hai//om

Ethnography for linguists

It is one of the innovative aspects of the DoBeS programme that the Volkswagen Foundation requires project teams to involve anthropologists as well as

towards decisions made by speakers “on their turf” in their specific situations, their positions and their locations. The result is that data sessions do not form a fixed mosaic of non-overlapping units but may be cross-cutting the pre-conceived categories of analysis of individual researchers.

Take the Hai//om healing dance as an example. On the basis of ethnographic work one may define wide sessions that are larger than a single video-tape and in fact larger than the dancing behaviour itself because it includes preparations and subsequent talk. Within this broad session ethnography also allows to define narrow sessions. These can be much shorter than the performance of a dance when they focus on the micro-interaction between healer and patients. The healer “rubbing the chest” of participants, or the healer applying scented powder to participants may be examples of such narrow sessions. Both types of sessions, broad and narrow, are grounded on the “turf” of the speakers instead of on pre-conceived ideas as to what constitutes “the healing dance” or the elements within it.

When defining data sessions, researchers can follow the typical course of ethnographic fieldwork. In ethnography we look out for how speakers distinguish elements on their own turf and we begin with a wide definition of units that are then funnelled down as our understanding increases. In other words: *Cast a wide ethnographic net and follow a funnel approach by defining wide sessions that can be narrowed down or broken down further in the course of research.*

What counts as ethnographic material?

It will have become clear from the discussion of the first question that ethnography is not a separate domain in the record (such as botany or technology) but it is the stuff that connects materials in different domains. It helps to gear language documentation towards a holistic perspective. One of the advantages of this holistic perspective is that relevant, and at times unexpected, connections appear. For instance the researcher may come to realize that kin-talk may involve plants or place names or the dead, cutting across pre-conceived domains. However, if ethnography is not simply to be added on top of a traditional linguistic description but in a sense saturates the language documentation as a whole, how do we manage this comprehensive whole and the many potential interconnections within it? In other words, if the body of materials gains a certain comprehensiveness (as is the aim in the DoBeS documentation and similar research programmes) the question of how the sessions are connected becomes important. If everything is *potentially* connected it becomes even more important to be able to trace such connections. Since ethnographic material relates to the process of involvement with the community of speakers in field research as well as to the social relationships that saturate cultural acts and objects it can provide a lead for establishing these connections.

The links between sessions that are highlighted by ethnography are based on the fact that it allows us to see specific speakers, places and institutions that are related to one another beyond the single speech event. In other words, any language documentation that makes use of ethnography will have in-built connections between sessions because the speakers (and addressees as well as bystanding listeners), the places and settings (spatial ones and non-spatial ones) and the social institutions are part of an ongoing process of mutual engagement which we tap when working ethnographically.

Take the Hai//om kinship system as an example. It is possible to systematically elicit a Hai//om kinship categorization by collecting genealogical data and by then asking informants as to what kinship term they would use to refer to the various individuals in this genealogical chart. Using ethnography as a research strategy it is possible to trace the relations of this informant (and others) not only on an abstract chart but as it is constituted across time and space. It will emerge that not only is there a difference between the kinship system of terms of reference and that of kinship terms of address used in practice. Differences also occur in the ways in which relatedness is being managed through diversified kin-talk. In small groups, in which everyone is related somehow genealogically, the idiom of kinship is much more than a label for a certain position in a genealogical chart. In a sense the kinship system is spread across a host of situations (economic, political, ritual) in which people refer to one another not necessarily in kinship terms but in what may be called “para-kin terms”, i.e. terms that define rights, demands and obligations of kin as an idiom of relatedness.

When looking for the “glue” that connects sessions, ethnography counts because it is more than just another domain. Ethnography helps to re-connect what has been archived as separate sessions. The connections are already there because unlike in anonymous survey-data or in large regional comparisons a collection of ethnographic data keeps the continuity of speakers (persons), places and settings across individual speech events on record. This fact can be exploited for the purpose of connecting sessions. In other words: *Do not elicit an abstraction of “cultural systems” (be it the kin categorization or any other system for that matter) in order to then subsequently search for its application but begin by collecting the pragmatics of (kin-)talk that allows to re-connect chunks of data and which still allows systematizations later on.*

How to locate ethnography in the record?

Defining sessions and being able to connect these records to a holistic view are problems that researchers face when compiling a language documentation archive. The problem as to how the data is to be accessed after it has been archived may seem less pressing at first but nevertheless has to be resolved from the start: If the

data sets are to be accessible by other researchers and across individual projects some standardized tools for data documentation need to be introduced (such as the IMDI editor and browser in the DoBeS programme). Metadata is an essential requirement from an archivist point of view but it is also an obvious place where ethnographic sources can be connected to the database. This applies to both, pre-existing ethnographic sources as well as materials collected in the project. In some cases the ethnographic material can be directly tied to the metadata categories if some flexibility is allowed for.

Metadata categories should - from the archivists' perspective - be standardized to some extent across the projects. However, this does not necessarily exclude the productive and - to some extent - also creative use of these categories by the individual researcher (or the individual research team). Surely, some free-text search function is - from the researchers' perspective - essential, allowing not only for the idiosyncrasies of each case but also for unforeseen changes in matters to do with the metadata. But even within the confines of the metadata templates used across cases there are opportunities for filling the categories with ethnographic life.

Take Hai//om spatial categories as an example. Spatial categories, just like categories of group or person, emerge at a number of levels and should therefore receive more than one entry in the metadata descriptions. In this case they would comprise (at least) toponyms, landscape terms and deictic language. Toponyms are manifold because there are several ethnic or language groups in one area and there is a dynamic history of settlement which is reflected in the naming of places. To note down the places of a recording (or of places referred to in conversation) is therefore a glimpse into a complex history of connected settlements that ethnography will bring to the surface. In Hai//om landscape terms are ubiquitous because they also feature in spatial language as absolute space markers. They connect people, landscape features, vegetation, social relations and directions and constitute one of the features that connect sessions that otherwise may be grouped separately as pointed out above. The use of deictic language and of gestures similarly involve both social and spatial elements and shorthand descriptions may occur in the metadata category of place as much as in the category of persons/groups involved.

When administering a growing body of data ethnographic input can balance the need for standardization. Metadata categories may seem to be a necessary evil but the fact that anyone using the archive in the future will encounter the metadata first should be incentive enough to "frontload" the metadata descriptions to the extent that outstanding ethnographic features should not be tucked away in remote folders but should be presented through a productive use of metadata descriptions. In other words:

Do not be judgmental about the diversity of space/time/person/group categories but make use of this diversity to facilitate access to the database.

There is no doubt that every experienced ethnographer will have more insights to convey which help to provide a better ethnographic record. The points raised in this short contribution are limited specifically to show how ethnography is a necessary and potentially rich feature in the work of interdisciplinary research teams working in the field of language documentation. There are other important entailments with regard to the relationship between researchers and the community of speakers which I have not touched upon but which deserve equal attention.