Chapter 7

Dictionary Day: A community-driven approach to dictionary compilation

Bryan D. Gelles
University of Florida

A common component of language documentation is the compilation of a small dictionary. The method of compilation has changed very little in the last century: most documentarians elicit individual lexical items from a speaker and check the item through both translation and backtranslation with other speakers. Two major problems with this method are the absence of larger community engagement and idiosyncratic problems that come from lexical item elicitation.

Animere is an endangered language spoken by around thirty speakers all aged over forty years. The speech community is located in Kecheibi, northern Volta Region, Ghana. Over a five month period I began the initial documentation of Animere with funds provided by a Small Grant from the Endangered Languages Documentation Programme, integrating Dictionary Day, one day a week when members of the community would gather to discuss lexical items. This method proved highly successful: I saved time and funds by making use of the speech community's intuition while obtaining valuable folk linguistic information when there was disagreement. Furthermore, the speech community was not only engaged but agentive, allowing for genuine consultation between the linguist and the speech community. The major drawback, however, is lack of synergy among documentarians and other linguists when excluding prescribed data collection methods.

1 Introduction

From the time of the Structuralists to the present, a language documentation at minimum consists of the Boasian trilogy: a grammar, a collection of texts, and a dictionary. The method for eliciting lexical items for a dictionary has also not changed much since the introduction of the Swadesh list (Swadesh 1955): most frequently a single linguist, using a Swadesh list, will elicit individual lexical items from a single speaker. As noted by Chelliah & de Reuse (2011) the dictionary has often been ignored by field linguists possibly due to limitations on time in the field and the linguist's research interests. Modern linguistic field methods and language documentation handbooks, however, devote

Bryan D. Gelles. Dictionary Day: A community-driven approach to dictionary compilation. In Jason Kandybowicz, Travis Major, Harold Torrence & Philip T. Duncan (eds.), *African linguistics on the prairie: Selected papers from the 45th Annual Conference on African Linguistics*, 105–116. Berlin: Language Science Press. DOI:10.5281/zenodo.1251720

an entire chapter on doing ethical fieldwork, focusing on collaboration with the speech community as well as encouraging linguists to give back to the speech community whenever possible, most frequently in the form of a tangible item such as a sketch grammar or dictionary.[1] During my own research, the first thing the speech community asked for was a dictionary of the language I was documenting. The compilation of a dictionary, regardless of its exhaustiveness, is one of the most straightforward ways of giving the community a physical manifestation of a documentation. Furthermore, Hill (2012) recounts a tumultuous field setting where the creation of a dictionary provided positive benefits to the speech community in the form of recognition of the language and the building of self-esteem in the community. A dictionary, thus, is not only a book of definitions but is a cultural icon for the speech community.

The question, however, is whether the method of compilation affects the type of lexical data is collected, whether for crosslinguistic comparison or theoretical application. If so, what methodological approaches should be used in order to strike a balance between the linguist's (and the larger linguistic community's) own goals and the desire to conduct an ethical documentation. Furthermore, the linguist must consider whether the status of the language affects what methodological approach should be taken. First I will discuss the current methods employed in the field for dictionary compilation as well as the implications for the wider linguistic community and the problems that come with it. Then I will present an alternative method I used when compiling a dictionary of a highly endangered language spoken in rural Ghana as well as discussing its implications and drawbacks. Finally I will offer concluding remarks.[2]

2 Current methodology

The current methodology for eliciting lexical items for a dictionary is mostly the same across fieldwork guides. Since Swadesh (1955), common practice has been for the linguist to elicit individual lexical items from members of the speech community. At one end of the extreme, Vaux et al. (2007) advocates for the lexical items to either come from a Swadesh list or a frequency list from a related language.

I	this	what	many	big
you	that	not	one	long
we	who	all	two	small

Figure 1: Sample portion of a Swadesh list

Mosel (2004), however, notes that a predetermined list such as the Swadesh list may present problems such as the absence of the lexical item in the speech community. For this reason, Mosel (2004) advocates for 'active eliciting' whereby the linguist asks for

[1] Among others Bowern (2008); Chelliah & de Reuse (2011), and Vaux et al. (2007).

[2] It should be noted the discussion itself is limited to lexical item elicitation and not specifically lexicography. A good summary of lexicography's own idiosyncratic problems can be found in Haviland (2006).

words related to a topic chosen by the linguist (for example, items in the home). Bowern (2008) also suggests allowing the community members to have limited agency by having the linguist ask them to do things like show them around the house and name items. The linguist, in all situations, records the lexical item for the dictionary. The linguist then must confirm the data with other speakers of the language through translation and backtranslation to account for inconsistencies throughout the speech community as well as to account for mistakes made on both the part of the linguist and individual community members. The data is then compiled into a dictionary.

Another way of eliciting data is by using field guides such as those found at the Max Planck Institute for Evolutionary Anthropology.[3] These stimulus kits consist mainly of pictures that the linguist points to, hopefully eliciting a lexical response as well as questionnaires. The linguist then records the data and uses it to compile a dictionary.

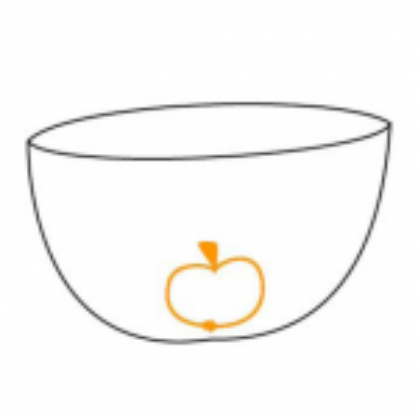

Figure 2: Example of topological relations stimuli (Bowerman & Pederson 1992).

Current language documentation methodology focuses on the audiovisual documentation of interactions with the speech community. At minimum, a microphone should capture the audio signal of both the speaker and the linguist to not only account for the lexical item but also the prompt (an invaluable resource when there are discrepancies among speakers). There should also be a video recording of the interaction to account for visual cues that may aid in spontaneous elicitations (the linguist may forget what exactly a given stimulus was for a lexical item whereas a recording will not). Dense metadata is compiled for each speaker and each interaction to account for not only foreseen circumstances (possible age, gender, dialect distinctions, etc.) but also for unforeseen circumstances (anything the linguist is currently unaware of about the community that may eventually play a factor in language differences across the community). The audiovisual component may capture those things that the linguist may either miss or misconstrue while gathering lexical items.

[3]http://www.eva.mpg.de/lingua/tools-at-lingboard/tools.php

In all of these instances, the linguist supplies the theme if not the individual items themselves, and (hopefully) the speech community provides a lexical item. After translating and backtranslating across multiple speakers, the linguist records what he or she believes is the consensus across the speech community. This data is compiled into a dictionary of the language with the necessary caveats in place regarding exhaustiveness.

3 Implications

It is not a coincidence that most linguists rely on a Swadesh list for gathering lexical items. The use of cognates to establish the genetic relationship between languages using the Comparative Method predates the Swadesh list, but the advent of the Swadesh list made this work much easier by codifying the list of words used for elicitation: since linguists were using (roughly) the same set of words in the field, typologists could use the data collected for direct comparison between languages. Also, by focusing on such things as numbers and color terms in a language, typologists are able to compare across multiple languages and language families relying on field linguists to gather this data during individual documentations. It is not uncommon for typologists to contact field linguists in order to see whether their documentations have such data necessary for typological work, an example of the synergy between typologist and field linguist that only a shared resource such as a Swadesh list can provide.

Stimulus materials are also used for crosslinguistic analysis by relying on the field linguist to gather very specific data the typologists and theorists cannot gather themselves due to logistical constraints. One example of this is the Pear Story, a student-made film that has grown to become a resource for analyzing crosslinguistic strategies for storytelling (Chafe 1980). Much like the Swadesh list, by having multiple field linguists use the same stimulus materials, typologists and theorists can analyze a specific type of data across languages without having to enter each speech community individually themselves, saving both time and finite resources.

4 Problems

There are several problems with these methods, however. As noted in Mosel (2004), the Swadesh list may not line up isomorphically with the language being discussed leading to inconsistencies: not only is it possible for an item on the Swadesh list to not be specific or general enough for the language in question (for example, a language that does not distinguish between the hand or the arm of a person), but it is possible that the item does not have a correlate in the language leading to an embarrassed speech community member who feels he or she is not up to the task at hand.[4] Though the former seems like a straightforward situation that will be easily noted by the linguist, other subtler lexical distinctions could be lost due to strict adherence to a predetermined list. The Swadesh

[4]During my own research, one community member, after not knowing a lexical item, checked every word with a family member during the rest of the session.

List (or a frequency list) thus may inherently fail to capture the semantic boundaries of the language while also possibly discouraging community members in the process.

The most problematic result of both using lists as well as 'active eliciting' is the difficulty in capturing cultural patterns due to working with individual speakers. The most common issue facing field linguists wishing to elicit lexical items is what to do with items that speakers disagree on. It is common in the field for one speaker to state that the lexical item is one thing, whereas a different speaker will insist that the first speaker has no idea what they are talking about and that the 'real' lexical item is something else entirely. In many cases, this may simply be due to dialectal differences between speakers, but if the linguist is unaware of such differences, this generalization may be lost and simply reduced to one speaker being incorrect. Furthermore, several speakers may differ from other speakers of the language. If the linguist discovers such a difference exists, he or she may deduce a generalization exists, but if the linguist only encounters a few members of the speech community and they all agree due to a small random sampling, this generalization is lost. In short, by only eliciting, translating, and backtranslating one speaker at a time, the linguist must assume the few speakers that were consulted were prototypical of the entire speech community, a flawed assumption statistically.

Another major problem with the Swadesh list specifically is how to elicit the items themselves. As can be seen in Figure 1 above, the Swadesh contains not only flora and fauna (which again may not exist in field site) but also items such as personal pronouns. Although the Swadesh list is suggested by field manuals, there is no explanation for how to actually elicit these items. As any linguist who has ever tried to elicit personal pronouns can attest, such lexical items are tricky at best to elicit. Furthermore, without any practical elicitation strategy to work from, each linguist creates their own method for elicitation often having to learn by trial and error what works and potentially misconstruing the data in the process. Although the list itself is codified, the way for eliciting it is not potentially leading to mistakes on the part of the linguist.

The linguist has also taken the majority of the agency of the documentation. The speech community, at best, has a choice among prescribed topics and at worst must merely translate from a list the linguist chooses. In this way the speaker is no longer a consultant who works with the linguist to document the language but is merely an informant who does the linguist's bidding. If the documentation is indeed a collaborative effort (a major emphasis from an ethical standpoint), it is disconcerting that the linguist is making decisions without the speech community's input in regards as to what lexical items comprise the language's dictionary. Literally, the linguist is telling the speech community what is appropriate for a dictionary that the linguist is partially using as justification that he or she is giving back to the community. In this power dynamic, the linguist has all the power, and the speech community is merely a group that has the data the linguist wants.

From a practical position, working with individual speakers is also a waste of time and resources. There is limited time in the field, and the linguist must manage this time wisely in order to accomplish as much possible. Working with individual speakers and then translating and backtranslating across individual speakers uses up not only time

but the resources the linguist is allocating for working with speakers of the community (compensation in whatever form the linguist deems appropriate). These resources could be used for other things that further the documentation as a whole instead of using primarily for the gathering of lexical items.

For these reasons, the current methods for gathering lexical items are insufficient. There is, however, another way to gather lexical items in a way that emphasizes collaboration while making differences among speakers clearer to the linguist. The question, though, is whether too much is then lost in terms of synergistic activities with typological and theoretical linguists.

5 Dictionary Day

From December 2012 until May 2013 I began the initial documentation of Animere, a Kwa language spoken in the rural northern Volta Region of Ghana. Previous contact with the community produced a sociolinguistic profile as well as a short wordlist used for comparative purposes (Ring 2006). The language is highly endangered, numbering around thirty speakers in one isolated village. The community consisted of cocoa farmers who work every day except for one day a week when the local market was held. On the morning of this day, all of the speech community was invited to participate in 'Dictionary Day',[5] a two hour period to discuss lexical items before they went to the market. Since the community determined it wanted a dictionary of their language, we agreed that I would use my linguistic resources to transcribe those items they deemed appropriate for their dictionary. They would decide on a topic for the day (or I would suggest a topic if they were at a loss for where to begin), and I would transcribe what they told me was appropriate for their dictionary. As this is a moribund language, it was common for the children of the speakers to come and watch the commotion, since Dictionary Day had a tendency to become rather lively at times due to disagreements. The dictionary that is being compiled of this language is organized based on the topics that the community (and sometimes myself) chose, including flaura, fauna, and traditional occupations. At the suggestion of one speaker, their dictionary includes useful phrases in the language as well. The dictionary, thus, is mostly their own work with the linguist performing the role of linguistic consultant as opposed to the guider of the elicitation.

6 Methodology

As opposed to the other methods for gathering lexical items, Dictionary Day is an attempt to gather the entire speech community at one time.[6] For reasons that are obvious this is not feasible in most field situations but is fully possible when the entire speech

[5]The name and the basic idea was first suggested by Dr. Jack Martin based on his collaboration with an American Indian speech community.

[6]It should be mentioned that Bowern (2008) states in passing that working in small groups was beneficial for collaborative reasons I will also mention for a larger group setting.

community is both small and local to the field site, and as will be seen this presents unique benefits that cannot be gained in much larger speech communities. The speech community is arranged in a circle, allowing each member full access to the conversation. The linguist is also a part of this circle as both physically and symbolically an equal part of the collaboration.

Figure 3: Dictionary Day

If the speech community meets one day a week, they are given the entire week to think about and discuss among themselves what they would like to be a part of their dictionary. By the time the linguist arrives on Dictionary Day, the topic will usually be selected already by the community. If this is not the case, the linguist can suggest topics that are appropriate to the speech community, allowing the community to determine whether they would like to proceed with the topic suggested.

Once a topic is suggested, the community members are asked to spontaneously suggest items for the dictionary. This will only have to be done once: the community will not need much prodding to suggest items in the future. With each topic the community members will discuss among themselves not only the appropriateness of the lexical item but also what forms to include in the dictionary. The linguist will then transcribe this form and use their linguistic expertise to identify relevant information about it for the sake of the dictionary, fulfilling their prescribed role of linguistic consultant.

An additional list of lexical items should be kept by the linguist with the community's permission. In this dictionary will be all the items that were controversial, noting the controversy surrounding the item and later, with help of the audiovisual record, what led to the disagreement. It will be this information that will shed light on the folklinguistics of the speech community as will be discussed further below.

All sessions should be recorded audio-visually, preferably from at least two angles if possible to capture all the community members. The audio component will rely on microphones with wide ranges in order to capture the spontaneous speech of the community members. For this reason microphone stands are essential: not only will the community start to speak more spontaneously without the constant reminder of a microphone that a linguist pointing at them would entail, but it is also impossible for a linguist to use a single microphone to capture all of the spontaneous interactions of the speech community. Although the linguist will be transcribing the dictionary on the spot and writing dense metadata about the session, it is these recordings that will reveal some of the missing cultural information the speaker does not know about the language ecology of the field site as will be explained below.

7 Implications

From an ethical standpoint, this method is ideal. The problem with the other methodologies is that they rely on the linguist to make all of the important decisions regarding what will go into the dictionary. As discussed above, if the linguist uses a predetermined list, the dictionary in effect becomes his or her work with the speech community only serving as informants rather than consultants of the project. Since current ethical guidelines call for a collaborative effort, the collaboration should not only extend to working *with* community members but also where possible to essentially work *for* them as well. It is worth stating that the majority of a language documentation has traditionally been to the benefit of the linguist as opposed to the speech community. This is one small way that the community itself is able to direct the documentation of their own language.

From a purely linguistic perspective, this method also alleviates most of the problems of the aforementioned methods. The question of how to elicit lexical items thus becomes moot. Instead of wondering how to elicit such items from the Swadesh list as 'louse' or 'I', the speech community will suggest items, negating any need for the linguist to invent idiosyncratic ways to elicit lexical items. Also, the problem of speakers not knowing a lexical item is no longer relevant as well. As Bowern (2008) notes, having multiple speakers during a session is beneficial in that speakers will be able to prompt each other on certain items that are little known among the speech community. This will alleviate the pressure on the speakers to perform for the linguist and will instead merely require the speaker to speak when comfortable, thus not endangering the linguist's relationship with individual speakers.

Another added benefit of this method is that disagreements among speakers are no longer in the hands of the linguist. As mentioned above, navigating discrepancies among speakers using a prescribed list falls on the linguist, since the linguist is meeting speakers

one at a time. Thus, if one speaker disagrees with another, the linguist must determine which speaker's item is suitable for the dictionary. This could cause a rift between the linguist and those speakers' items that were left out of the dictionary, since it is the linguist who determines the veracity of each item. If the decision is left to the community, this no longer becomes a problem. From a practical standpoint, it is also incumbent on the linguist working with individual speakers to determine what constitutes a representative sample. Field manuals mention translating and backtranslating as a way of policing data, but they fail to mention just how many times it is required before an item is acceptable to add to the dictionary, leaving the choice to each individual linguist. Such an unsystematic approach could lead to idiosyncratic data, a situation often found when dealing with older language data. This linguistic policing of data is no longer the job of the linguist but falls onto the speech community, the group that has a better knowledge of the language and the idiosyncrasies that come with it.

Disagreements, however, are also important for linguistic information that is normally unavailable to a linguist working with a new speech community one member at a time. Through disagreements among the speech community, the linguist can glean sociolinguistic information about the language. During a heated debate during Dictionary Day, two groups formed, arguing about which lexical item was most appropriate for the language. Both sides claimed the other was wrong, and neither was willing to give any ground. Through mediation among other members of the speech community, a form was selected for their dictionary. My dictionary of the language, however, has both, because the two groups that were arguing belonged to different age groups: the age-mates of one group were arguing with the age-mates of the other. Though currently unprovable, this suggests that there may be a generational difference linguistically that I may have not seen if I had approached each member one at a time. During another session, the leader of the speech community suggested an item, and everyone automatically supported the item due to the speaker's prestige. One speaker, however, disagreed, telling me privately that another form was preferable. This form turned out to be an extension of a morphological pattern that I had not seen previously. Without this quiet reaction from a member normally not vocal, I would not have seen the pattern. In this way, through various spontaneous disagreements over otherwise uninteresting lexical items, I was able to discover both sociolinguistic data as well as a linguistic pattern I would not have been able to see previously.

One major benefit of Dictionary Day that has thus far been assumed is the idea of consensus among speakers. Using traditional methods, consensus is a matter of the linguist determining just how many members are necessary to constitute a representation of the entire language. When working with a small speech community, this can be done by speaking to each community member individually, but, as mentioned, disagreements must be navigated somehow by the linguist. By bringing the entire community together, however, consensus can be built among the community itself. By discussing items individually among themselves, they are literally forming a consensus for each item one by one. Verification is done on the spot without any need to recheck most items individually. When a major dispute occurs, however, it becomes necessary to approach in-

dividual members of the community to determine what constitutes speaker differences. This, however, is only limited to major disputes, whereas the traditional method requires rechecking every item. In short, actual consensus among the community can be reached by having the entire community present at one time as opposed to choosing a number of speakers to individually confirm lexical items.

Finally, from a practical standpoint, Dictionary Day saves both time and resources. Instead of having to allocate the beginning of each day to checking and rechecking various lexical items speaker by speaker, the linguist can use one day a week to go over the same amount of words while freeing up the rest of the week to work on other things. Since each lexical item is verified at the time of its suggestion, no additional time is required, and more lexical items can be elicited quickly and efficiently. Also, whatever the linguist deems appropriate in terms of compensation to the community will be used towards other things besides gathering lexical items, a boon to the linguist who may have personal goals in mind in the field.

Dictionary Day thus solves the problems presented by the traditional method of gathering lexical items. Through real collaboration with the speech community, the linguist is not only ethically interacting with the community but also doing it in a way that that benefits his or her own research goals by freeing up additional resources. More importantly, the idiosyncrasies of the data can be worked out in a group setting without the linguist becoming the arbitrator. The linguist may also discover language patterns that would not be visible when speaking to only a single member of the community, a help to the field linguist who is documenting a language that has not been analyzed previously.

8 Problems

When compared to methods that require the linguist to choose topics that speakers then supply lexical items for, Dictionary Day is preferable in all respects. However, when compared to the use of prescribed lists or stimulus kits, Dictionary Day has a major drawback, namely synergy among theorists, typologists, and field linguists. As previously mentioned, by using a Swadesh list, field linguists are supplying comparative linguists with data that they themselves cannot obtain. Also, by using stimulus kits, the field linguist is no longer supplying theorists and typologists with the same kind of crosslinguistic data. Although language documentation is itself becoming an independent field with its own goals, it is still preferable for documentarians to work with other linguists rather than isolate themselves in their subfield. A common refrain among documentarians is that it is not their job to orient their documentation around prescribed data collection methods by theorists and typologists. It is also a common refrain among documentarians of understudied languages that their work is often ignored by those same theorists and typologists that they themselves refuse to work with. By building a documentation collaboratively with the speech community, the data gathered becomes idiosyncratic in that it may not fulfill any needs of other linguists due to the random sampling of data in the field. In this way, Dictionary Day further exacerbates this problem by not only not limiting the data to prescribed areas of interest to other linguists but also by possibly failing to address such areas at all.

It is, however, worth noting that many field linguists choose to use a Swadesh list not due to any concern with other linguists' interests but due to not contemplating an alternative, and it is very common for fieldwork to go unnoticed regardless of the field linguist's intentions to the contrary. These are much larger problems than one methodology could possibly address, but it is worth mentioning those areas where the methodology fails to bridge the gap between documentarians and other linguists. For this reason, Dictionary Day should be used in collaboration with more traditional methods. A simple way of addressing this issue is to add Swadesh list items whenever possible to Dictionary Day itself when the community allows. Stimulus kits could also be added, though practically it seems out of place in the context of lexical item elicitation. Whenever possible, both traditional methods and Dictionary Day should be used side-by-side in order to not only address the problems of the former but also to account for the problems with latter. In this way, the documentarian can work with other linguists while not compromising the collaborative goals of the documentation.

9 Conclusion

Dictionary Day is a way for a field linguist to work collaboratively with a speech community as a whole in situations where such a collaboration is feasible. Considering the concern of documentarians with the ethics of fieldwork, such a speech community-driven collaboration is preferable, since it gives the agency to the community as opposed to the linguist who has traditionally not only had all of the power but mostly uses such power to guide the documentation in the direction of his or her own research interests. Although direct elicitation is making a comeback (Matthewson 2004), allowing speakers to spontaneously suggest lexical items reduces the problems of elicitation such as data reliability. It also benefits the data collection by not only offering a different mechanism for dealing with disputes among community members but also using such moments to gain insights into the language itself. Consensus is thus built among the entire speech community and not left to the linguist to determine what arbitrary number constitutes speech community consensus. Practically, it also saves time in the field for furthering the documentation in other ways while the linguist is in the field.

Problematically, though, Dictionary Day fails to account for linguists who need cross-linguistic data. By focusing solely on what the community chooses to do, the field linguist is not feeding more new and interesting data into the comparative, theoretical, and typological discussion that a Swadesh list or stimulus kit would. For this reason, Dictionary Day should be used in collaboration with other methods whenever possible. The community's wishes must come first, but the linguist still has an obligation to the field if he or she hopes to address such issues as the absence of understudied languages in linguistic theory. Although documentarians and other linguists sometimes have disputes about the exhaustiveness of linguistic typology and theory, the impetus is on the documentarian to enter the discussion as well. By combining both methods, the field linguist can find a way to bridge the divide between documentation and theory.

Acknowledgements

This research was made possible by a grant from the Endangered Languages Documentation Programme (SG0199).

References

Bowerman, Melissa & Eric Pederson. 1992. Topological relations picture series. In Stephen C. Levinson (ed.), *Space stimuli kit 1.2*, 51. Nijmegen: Max Planck Institute for Psycholinguistics.

Bowern, Claire. 2008. *Linguistic fieldwork: A practical guide*. New York: Palgrave.

Chafe, Wallace L. (ed.). 1980. *The pear stories: Cognitive, cultural, and linguistic aspect of narrative production (Advances in discourse processes)*. Santa Barbara: Praeger.

Chelliah, Shobhana & Willem de Reuse. 2011. *Handbook of descriptive linguistic fieldwork*. New York: Springer.

Haviland, John Beard. 2006. Documenting lexical knowledge. In Jost Gippert, Nikolaus P. Himmelmann & Ulrike Mosel (eds.), *Essentials of language documentation*, 129–162. Berlin: Mouton de Gruyter.

Hill, Deborah. 2012. One community's post-conflict response to a dictionary project. *Language Documentation and Conservation* 6. 273–281.

Matthewson, Lisa. 2004. On the methodology of semantic fieldwork. *International Journal of American Linguistics* 70. 369–415.

Mosel, Ulrike. 2004. Dictionary making in endangered speech communities. In Peter Austin (ed.), *Language documentation and description, Endangered Languages Project*, 39–54. London: School of Oriental & African Studies.

Ring, J. Andrew. 2006. *We have no one to sing our songs: Concerns of an African elder*. Presented at the International Workshop on the Documentation and Description of GTM Languages.

Swadesh, Morris. 1955. Towards greater accuracy in lexicostatistic dating. *International Journal of American Linguistics* 21. 121–137.

Vaux, Bert, Justin Cooper & Emily Tucker. 2007. *Linguistic field methods*. Eugene: Wipf & Stock.