Is ex situ documentation valid? Language documentation in immigrant and refugee communities

Brent Henderson

Proceedings of Conference on

Language Documentation & Linguistic Theory 3

Edited by Peter K. Austin, Oliver Bond, Lutz Marten & David Nathan

19-20 November 2011 School of Oriental and African Studies, University of London

Hans Rausing Endangered Languages Project
Department of Linguistics
School of Oriental and African Studies
Thornhaugh Street, Russell Square
London WC1H 0XG
United Kingdom

Department of Linguistics:
Tel: +44-20-7898-4640
Fax: +44-20-7898-4679
linguistics@soas.ac.uk
http://www.soas.ac.uk/academics/departments/linguistics

Hans Rausing Endangered Languages Project:
Tel: +44-20-7898-4640
Fax. +44-20-7898-4349
elap@soas.ac.uk
http://www.hrelp.org

© 2011 Brent Henderson

No part of this publication may be reproduced, stored in a retrieval system, or transmitted, on any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the author(s) of that part of the publication, except as permitted by UK copyright law.

ISBN: 978-0-7286-0398-1

This publication can be cited as:

Brent Henderson. 2011. Is ex situ documentation valid? Language documentation in immigrant and refugee communities. In Peter K. Austin, Oliver Bond, Lutz Marten & David Nathan (eds) *Proceedings of Conference on Language Documentation and Linguistic Theory 3*, 131-140. London: SOAS.

or:

Brent Henderson. 2011. Is ex situ documentation valid? Language documentation in immigrant and refugee communities. In Peter K. Austin, Oliver Bond, Lutz Marten & David Nathan (eds) *Proceedings of Conference on Language Documentation and Linguistic Theory 3*. London: SOAS. www.hrelp.org/eprints/ldlt3_14.pdf

Is ex situ documentation valid?
Language documentation in immigrant and refugee communities

BRENT HENDERSON
University of Florida

1. INTRODUCTION

As an enterprise, language documentation (LD) is defined by a general set of goals aimed at creating 'a lasting, multipurpose record of a language.' (Himmelmann 2006: 1). However, it is recognized that LD methodologies must be tailored to particular documentation efforts. This paper discusses issues that surround doing 'ex situ LD' in the context of immigrant and refugee communities, a small but growing practice in the field. I will argue that, despite objections raised in the literature on linguistics field methods against working in such communities, ex situ LD is not only valid, failing to differ from 'in situ' LD in any qualitative way, but is also increasingly urgent, given global trends in shifting human geography.

The paper is organized as follows: in Section 2, I argue that the impact of community displacements on language loss has been underestimated, and that those interested in language documentation efforts cannot afford to ignore refugee and immigrant communities. In Section 3, I discuss common objections to ex situ LD, countering that the practice does not really differ qualitatively from more traditional efforts in its considerations. In Section 4, I discuss some practical methodological issues from the perspective of ex situ LD, based on two years of experience documenting the Somali Bantu language Chimwiini in Atlanta, London, and Mombasa.

2. IS EX-SITU LD NECESSARY?

In setting out a justification for ex situ LD, the first question that arises is whether such work is actually necessary. Of course, underlying that question itself is the assumption that ex situ LD should be dispreferred (a position challenged in the next section), but I put that aside for now. There are at least two situations where the necessity of ex situ LD is recognized, namely when the home location of a language cannot be visited by researchers due to safety concerns (such as political instability or the effects of famine), or when the remaining speakers of a language are so few the language they speak cannot be said to have a home any more.

The former situation is exemplified by Chiimwiini, a language I have worked to document for the past two years. Chimwiini is a language of a single town (Brava, or Miiini) on the southern Somali coast, and is often referred to as a dialect of Swahili. When the Somali Civil War began in the late 1980s, the town was ravaged more than once by militia forces and those who could find the means to flee to neighboring Kenya did so. At least half of all Chimwiini speakers were eventually given UN refugee status and relocated to the United States and the UK, while many remain in

Kenya where a slow but steady trickle of speakers continues to arrive from Somalia. The town of Brava, home to the language, continues to be occupied by militant forces who know nothing of the language or culture of the Wantu waMiini and have taken steps to actively destroy local cultural symbols (especially the tombs of venerated saints). Clearly in this situation ex situ LD is not only justified, but urgently necessary. Working with the Wantu waMiini communities in the US and UK, I have noted that most speakers under the age of 25 are semi-speakers, able to converse in the language but unfamiliar with finer distinctions in syntax or morphology. In Mombasa, Kenya, the language is somewhat more stable and even younger community members can speak the language well. Amongst themselves, however, younger speakers are as likely to speak Swahili as Chimwiini and there is significant loss of distinctive vocabulary. Given these facts, it is doubtful that Chimwiini will continue to be spoken fluently by the current youngest generation or, at the very least, that the Chimwiini this generation speaks will be radically influenced by English and Swahili at every level of structure.

Of course, the influence of war and community displacement on language shift has always been recognized. However, it has seldom been seen as a major impact of language loss. Rather, language death is largely seen as a gradual process resulting from sustained contact between socio-politically unequal languages (e.g., see Batibo 2005:63-65). The real impact of war on language loss is unknown, however, since to my knowledge research on this area is virtually non-existent. Perhaps this is for good reasons: in the midst of widespread death, starvation and disease, providing support for language preservation may register low on the priority list, as Brezinger (2009:38) has suggested. Nevertheless, given that it is well-known that large-scale community displacement leads to language loss, and given the statistics available to us regarding human displacement and refugee status, it is easy to surmise that the effects of war on language loss have been underestimated and may be on the rise.

According to the International Displacement Monitoring Centre[1], in 2010 there were more than 27 million internally displaced persons (IDPs) around the globe, along with 13-16 million refugees. While the annual number of refugees has been in this range for more than a decade, the number of IDPs has shown a general uptrend over the past decade, increasing from 17 million in 1997. Nearly half of current IDPs (over 12 million) are in Africa, many in linguistically diverse regions. In 2010, Somalia had 1.5 million IDPs due to its ongoing civil war involving most from the southern half of the country, its only linguistically diverse area. The disproportionate manner in which minority linguistic groups have been targeted in that conflict has led Nurse (2010) to declare the end of Bantu languages in Somalia. Minority Cushitic languages like Jiddu and Tunni may suffer similar fates.

The Democratic Republic of the Congo in 2010 had 1.9 million IDPs in its eastern provinces of North and South Kivu where more than a dozen language communities reside. Topping the list in Africa is the Sudan with nearly 5 million IDPs, most of whom are from the southern half of the country which is home to rich linguistic diversity. Ethnologue lists more than 40 languages in the Nuba Hills region alone, many with only a few thousand speakers each. The precise impact of these

[1] See www.international-displacement.org.

displacements and others on linguistic diversity is unknown, but must surely be severe. In many cases the displacement has lasted for extended periods of time. Sudan in particular offers an interesting case study. With more than 4 million linguistically diverse groups displaced for more than a decade, the recent relative peace and now independence of the south has led to the return of hundreds of thousands of IDPs to the new Republic of South Sudan. One wonders what the impact of the displacement and now return of so many speakers will be on linguistic diversity in the region, considering that the returning groups may include thousands of younger individuals who do not speak the language of their homelands as a first language. The shifts in language use dynamics are likely to be swift and complex.[2]

The question of ex situ LD, then, is not merely one of preference, but one of necessity in situations of community displacement, be it due to permanent exile of refugees or temporary exile as IDPs. My sense is that few would take issue with this argument. Where the issues might lie, then, is in choosing to conduct language documentation efforts in immigrant communities when non-immigrant communities might be more or less accessible. I turn to this in the next section.

3. IS EX-SITU LD DISPREFERRED?

There is a bipartite caricature that runs through many standardly-used manuals for linguistic fieldwork, perhaps most explicitly (and intentionally comically) spelled out by Crowley (2007:12) as a contrast between the 'arm chair' linguist who does their work 'without putting in any effort to conduct original research on any language' and the 'dirty feet' linguist who 'gets actively involved in the study of previously undescribed languages.' Crowley also admits there is a 'kind of halfway house' between these two extremes, namely doing fieldwork on a language in a local urban context. As Crowley (2007: 13) puts it:

> I have seen some academic curricula vitae in which linguists have stated that they have 'done fieldwork' on such-and-such a language, though in reality they may have travelled no further than the outer suburbs of San Francisco or Manchester…At most, this kind of fieldwork is useful if you are only interested in studying a particular feature of a language without intending to produce a coherent overall account.

Having actually conducted fieldwork in Manchester, I must slightly protest. What Crowley (perhaps unfairly singled out here from many other fieldwork manuals) fails

---

[2] As stated above, one might question the appropriateness of conducting linguistic fieldwork in the context of humanitarian crises; however, I do not think this should be ruled out in principle. Indeed, if one considers 'linguistic security' to be a human need, as I think many would, then basic survey and documentation work on local language dynamics can be argued to be an important part of humanitarian efforts. Add to this the insight that knowledge of local languages and cultures makes humanitarian work more efficient and effective and there is ample reason to suggest that linguistic work should be an integral part of such efforts.

to consider is that there might actually be a robust, cohesive community of speakers in Manchester's suburbs who use their language on an everyday basis, speak it alone in their homes with family, and may even include rather surprising numbers of monolingual speakers. I certainly found this to be the case in Manchester as well as in London. While younger Chimwiini speakers are certainly undergoing language shift to English, many speakers over age fifty that I interviewed spoke very little English, despite having been in the UK for fifteen years. Women with competence in English were fewer in number and many were monolingual in Chimwiini or bilingual in Chimwiini and a Somali dialect.

Strong shifts in human geography due to immigration and forced displacements (see Section 2 above) have made this situation a common one in urban centers around the globe. In most major cities there are sizeable, fairly monolithic enclaves of people practicing foreign cultures and speaking foreign languages. Yet, for the most part, such communities are not seen as potential partners for documentation efforts, despite the fact that working with these communities has significant advantages (but see Kaufman 2009 for a discussion of practical advantages of ex situ documentation).

One reason immigrant communities are not seen as partners for documentation efforts is that such languages are seen as undergoing rapid language shift and change. As Aikhenvald (2007:5) puts it:

> Working with immigrant communities – if a language is well spoken in a home country – is also hardly advisable: many grammatical features are extremely prone to contact induced change and are likely to shift under the impact of introducing new – and losing old – cultural practices. Opting to study Burmese, Hmong, or Serbian – each spoken in their home countries – within the comfort zone of Greater Melbourne, London or Los Angeles may be good for understanding the subtle influence of the Anglophone environment on a smaller language. But it is bound to give a skewed picture of the language's structure.

Such a perspective is that, given the option, one should invariably choose to conduct LD in the 'home country' rather than in immigrant communities. This clearly brings a greater challenge to the validity of ex situ LD than the caricatured dichotomy of linguists doing 'real fieldwork' and those who never leave the confines of a university office. The intuition behind Aikhenvald's preference is easy to sympathize with. Languages, like species, adapt to their environments, incorporating grammatical categories and vocabulary that are the most relevant for everyday life. When that everyday life experience and environment suddenly change radically, as in the immigrant situation, languages may adapt rapidly as well. Though a community's home language may continue to be the first language in the home and community interactions, the language of the broader society and culture is bound to have a strong influence, perhaps even, as Aikhenvald says, at the structural level.[3] While this may

[3] I am being extremely unfair in singling out Aikhenvald (2007) here; it just happens that this paper clearly expresses the bias against doing fieldwork in immigrant communities. It should not be seen as a negative judgment on that particular paper or author.

indeed be true, I would like to argue that it is not a reason for abandoning the idea of ex situ LD in immigrant communities. In fact, I propose that LD in an immigrant community does not differ in qualitative, fundamental ways from LD in a language's home country, though of course special considerations may apply (as they must in any specific LD endeavor). Below I address two potential concerns with this stance.

3.1. Contact-induced change

It is certainly true that immigrant or otherwise displaced language communities will experience pressures of language shift and contact-induced change as younger speakers grow up bilingual and older speakers may find they use the lingua franca of their new home more than their first language(s). There are two reasons, however, why this is not a valid reason for excluding ex situ LD. First, as Bowern (2008) has pointed out, language contact is often over-estimated in immigrant communities, which often turn out to be rather cohesive and even stable over time with regard to their language use. When traveling through eastern London or Atlanta, one is not necessarily struck by a sense of vast diversity, but rather by a sense of homogeneity as many neighborhoods have loose ethnic and linguistic boundaries. Each locale may have its own shops, restaurants and markets where one often finds business owners who speak the same languages as the local population. Often speakers who have immigrated as adults do not find an urgent need to use English in their everyday lives. In London, I met many elder Wantu waMiini and Somalis who have lived there for more than fifteen years, yet have not become conversant in English beyond basic greetings.

Second, contact-induced change is by no means unique to the immigrant community situation. Indeed, if one surveys the language contact situation of most minority languages in the world, instances of contact-induced change must be concluded to be the norm for small languages rather than the exception. The majority of minority languages are under pressure from larger regional lingua francas, languages in which minority language speakers are often bilingual. It is difficult to see, then, how the contact situation is qualitatively different for a community of Chimwiini speakers in Columbus, Ohio than it is for, say, Ndengeleko speakers in Tanzania who are under extreme pressure to use and shift to Swahili (see Ström 2009). Contact-induced change, then, is a factor in many LD contexts, regardless of whether such efforts are taking place in the home country or not.

3.2. Getting a skewed picture of the language

A second, related concern is that documenting a language outside its home country will lead one to a 'skewed picture' of the language. Clearly the underlying assumption is that the language as spoken in the home country is in some way more pure, uncontaminated, or otherwise legitimate for documentation efforts. Again, one can understand the intuition here: there are certainly rich environmental connections that a language has to the places it has been spoken for long periods of time that will not exist in a new environment.

Nevertheless, we should be suspicious of this bias, not only for the reasons mentioned in the previous section, but because we come dangerously close to falling

into the trap of believing that any LD effort ever really documents anything we commonly refer to as 'a language.' Though we often talk about languages as if they exist in isolation, languages are dynamic social systems made up of the collective consistencies and variation in the idiolects of the people who speak them. Therefore, all any LD effort ever does is document particular speakers in particular contexts. Of course, LD efforts attempt to be as comprehensive as possible, documenting a large variety of speakers in a large variety of contexts. Such efforts will always be incomplete, however, and will always involve decisions regarding which speakers and which contexts to pay attention to.

This is not only true synchronically, but diachronically as well. A thousand years ago, Chimwiini was part of a continuum of what were likely closely similar dialects of Swahili spoken all along the coast of East Africa from northern Mozambique to Mogadishu. In the late 15th century, Cushitic-speaking groups from the interior pushed to the coast in Somalia, gradually supplanting Swahili with Somali dialects in coastal towns and villages in most areas. The town of Brava and its Chimwiini-speaking population were spared, however, possibly due to early-formed political alliances with several Tunni and Somali groups who were already fully integrated into the life of the town, speaking Chimwiini as their second language. This led, however, to many Somali loanwords being borrowed into Chimwiini, affecting the language's vocabulary, phonology, and morphology. Later, Omani occupation of the Swahili coast, along with consistent waves of Arab immigrants from Oman and Yemen through the centuries, led to massive influxes of Arab loan words and possibly more structural influences as well (see Abasheikh 1976). In the 1970s government policy mandated that all citizens learn Somali, with the result that most Chimwiini speakers today are bilingual in Somali.

The point of this brief history is to illustrate that at no point in the history of Chimwiini was the language free from external pressures to undergo change. It is difficult to argue for one period or another as the focus for 'a documentation of Chimwiini.' LD efforts always make choices about the contexts and individuals they document. The question becomes, then, is there a reason speakers in an immigrant or refugee context should not be included as one of these contexts? And if, for some reason, one's set of documentation contexts is limited to immigrant contexts, is this a reason an LD effort should not go forward? I cannot think of a good argument against it.

4. PRACTICAL ADVICE FOR EX SITU LD

Below I offer some practical advice for undertaking LD in immigrant or refugee communities in major urban centers. The notes presented here are by no means comprehensive or even systematic, but represent my experiences conducting fieldwork on Chimwiini in various urban centers over the past two years.

4.1. Personal considerations

Any field worker must take personal and psychological sensitivities of the communities they work with into account. For refugee communities in particular, one must be aware that individual members as well as the community as a whole have

been through traumatic events that may have left long-lasting scars, making certain topics off-limits or very sensitive. On the other hand, one also finds that some community members are eager to tell the story of these experiences so they become better known to outsiders. Navigating these waters can be tough, especially if there are competing opinions about the community's political situation. For this reason I have tried to largely stick to historical narratives ('what happened?') and avoid more political discussion ('why did this happen?') when discussing recent historical events.

One must also keep in mind that often individuals in refugee or immigrant communities have come to their new homes via often unpleasant dealings with Western bureaucrats and officials, and some may have entered the country illegally. Suspicion of outsiders (particularly those asking them to sign consent forms while using a microphone!) can be quite high. Giving assurances about privacy and clearly explaining the nature of the work are therefore crucial, though in my experience many people are still simply not willing to take the risk that something they say could get them in trouble with the authorities. For similar reasons, many people may consent to be recorded via audio, but not video.

4.2. The urban environment

Generally people in immigrant communities are simply very busy. Many work more than one job or are very active in the 'informal economy.' This goes double for community members who are most likely to fully appreciate the benefits and goals of an LD endeavor, namely community leaders who, in addition to work and family, also act as informal social workers, helping with immigration paperwork, organizing community events, and even leading religious services. This makes it difficult to be picky about whom one works with and when. Trying to schedule appointments on the researcher's timeline is bound to fail. Rather, I have found it necessary to simply be available, ready to meet someone at a moment's notice when they have the time. This requires being as local to the community as possible. While working in Mombasa, I rented an office just across from the mosque which many Wantu waMiini men attend, making it clear that people could drop in to speak with me at any time. Of course, this approach often causes one to lose some control over one's recording environment. Combined with the fact that urban environments come with car engines, police sirens, and ubiquitous mobile phones, background noise is an issue and trying to minimize it is typically a losing battle. The best approach is simply to have good external microphones that can be directed at speakers and to accept that some background noise in recordings is unavoidable.

4.3. Technology

One of the great advantages of working with immigrant communities in urban centers (as opposed to rural locals in the developing world) is that many have good computer skills and regular access to technology and the internet. This makes communication during and after fieldwork easy and productive. Using the internet, we have developed a regular workflow between individuals on three continents for doing transcription, translation, and dictionary work. Access to technology also allows fieldwork to continue productively after the researcher has departed, as trained

assistants can continue to record interviews and send the digital recordings for analysis and archiving.[4]

4.4. Young semi-speakers

While not often a reliable source of primary data, young semi-speakers can be an invaluable resource in documenting a language in an immigrant context. These individuals are often fully fluent in both the language and culture of their urban context, yet still have key insights into their heritage language and background. They can therefore serve as a resource of cultural insight for the researcher, especially regarding delicate or sensitive issues that it may be more difficult to ask elders or community leaders about. Most importantly, however, younger community members are often uniquely motivated to assist in documentation efforts. Many are curious about the history and culture of their parents and have questions about their own cultural identities.[5] Because of this, some will jump at the chance to be involved in the LD effort by transcribing, translating, and conducting interviews. In addition, younger speakers are generally more technologically proficient, making them easy to connect with over the internet (see 4.3. above).

4.5. Topics to target

It is traditional in documentation efforts to target certain culturally-rich speech events (weddings, funerals, dedications, etc.) for recording and analysis. In this way, such efforts are very present-focused, seeking to document the language as it is used currently. In the immigrant context, however, such cultural practices will have changed radically, if they are still practiced at all. I have spoken at length with people about the details of Bravanese weddings, which are multi-day events with detailed prescribed roles for specific family and community members. The weddings I have attended in Atlanta, Mombasa, and London, however, have looked much closer to modern western weddings: simple ceremonies followed by a meal in a large wedding hall.[6] Because of these shifts in practices, LD in an immigrant context may find it more productive to be somewhat backward-focused and future-focused, rather than present-focused. By backward-focused, I mean recording narratives about the ways the language and culture used to be practiced before the exodus from the home country. Often one finds that this is a primary concern of the elders of the community who are keenly aware of the value of traditional knowledge and how quickly it is disappearing. By future-focused, I mean focusing on documenting linguistic

---

[4] I must admit I did not realize the full potential of this sort of collaboration until fairly recently. I am now convinced that envisioning LD as a network of researchers and language speakers actively collaborating online should be a first step in any LD endeavor for which the technology is available.

[5] This is certainly the case in the Chimwiini communities I have spent time in. Many young speakers are very concerned about questions of identity, clan affiliation, and origins, questions I am told almost never arose before the exodus from Somalia.

[6] Admittedly, I have only had access to the men's experience at these events. I am told the bride's wedding experience is somewhat more traditional, involving certain songs and dances.

knowledge that is most likely to be useful to future generations who may want to recapture their culture and language, even if they are not native speakers themselves. Here I am thinking of the kind of everyday language (greetings, introductions, and other basic conversational conventions) that is unlikely to have changed much in the transition from the home country. This set of linguistic knowledge is often overlooked in documentation efforts, perhaps because it seems so mundane. However, as Amery (2009) has argued, these aspects of language are often what future heritage speakers find most useful.

5. CONCLUSIONS

In this paper, I have highlighted the fact that high numbers of displaced peoples around the globe and the growth of urban centers as collectives of cultural enclaves mean that linguists should not ignore the possibility of doing documentation in displaced communities (be they immigrant, refugee, or IDP groups). I have also argued that such ex situ LD does not fundamentally differ in its goals or concerns from LD that takes place in the language speakers' home country; in addition, I have also offered some practical advice on how focuses and methodologies may differ.

REFERENCES

Abasheikh, Mohammed Imam. Reflexivatization in Chimwi:Ni. *Studies in the Linguistic Sciences* 6, 1-22.

Aikhenvald, Alexandra. 2007. Linguistic fieldwork: setting the scene. *Sprach-typologie und Universalienforschung* 60, 1-11.

Amery, Rob. 2009. Phoenix or relic? Documentation of languages with revitalization in mind. *Language Documentation & Conservation* 3, 138-148.

Batibo, Herman. 2005. *Language decline and death in Africa: causes, consequences, and challenges.* London: Multilingual Matters.

Bowern, Claire. 2010. Fieldwork in Language Contact Situations. In Ray Hickey (ed.), *The Handbook of Language Contact,* 340-358. Oxford: Blackwell.

Brezinger, Matthias. 2009. Language diversity and poverty in Africa. In Wayne Harbert, et al. (eds.), *Language and Poverty*, 37-49. London: Multilingual Matters.

Crowley, Terry. 2007. *Field Linguistics: A Beginner's Guide*. Nick Thieberger (ed.). Oxford: Oxford University Press.

Dimmendaal, Gerrit. 2001. Places and people: field sites and informants. In Paul Newman and Martha Ratliff (eds.), *Linguistic Fieldwork*, 55-75. Cambridge: Cambridge University Press.

Himmelman, Nikolaus. 2006. Language documentation: What is it and what is it good for? In Jost Gippert, Nikolaus Himmelmann & Ulrike Mosel (eds.), *Essentials of Language Documentation*, 1-30. Berlin: Mouton de Gruyter.

Kaufman, Daniel. 2009. Ex-situ language documentation and the Urbana Fieldstation for Linguistic Research. Talk presented at the 1st International Conference on Language Documentation and Conservation (ICLDC). Online at http://scholarspace.manoa.hawaii.edu/handle/10125/5065.

Nurse, Derek. 2010. In F. Floricic (ed.), *Essais de typologie et de linguistique générale. Mélanges offerts à Denis Creissels*, 187-200. Lyon. ENS Editions.

Ström, Eva-Marie. 2009. The situation of Ndengeleko, a coastal Tanzanian language (P10). In Masangu Matondo, Fiona McLaughlin, and Eric Potsdam (eds.), *Selected Proceedings of the 38th Annual Conference on African Linguistics*, 229-241. Somerville, MA: Cascadilla.