Endangered Languages and Languages in Danger

Issues of documentation, policy, and language rights

EDITED BY

Luna Filipović
Martin Pütz

Endangered Languages and Languages in Danger

IMPACT: *Studies in Language and Society*

ISSN 1385-7908

IMPACT publishes monographs, collective volumes, and text books on topics in sociolinguistics. The scope of the series is broad, with special emphasis on areas such as language planning and language policies; language conflict and language death; language standards and language change; dialectology; diglossia; discourse studies; language and social identity (gender, ethnicity, class, ideology); and history and methods of sociolinguistics.

For an overview of all books published in this series, please see
http://benjamins.com/catalog/impact

General Editors

Ana Deumert
University of Cape Town

Kristine Horner
University of Sheffield

Advisory Board

Peter Auer
University of Freiburg

Jan Blommaert
Ghent University

Annick De Houwer
University of Erfurt

J. Joseph Errington
Yale University

Anna Maria Escobar
University of Illinois at Urbana

Guus Extra
Tilburg University

Marlis Hellinger
University of Frankfurt am Main

Elizabeth Lanza
University of Oslo

William Labov
University of Pennsylvania

Peter L. Patrick
University of Essex

Jeanine Treffers-Daller
University of the West of England

Victor Webb
University of Pretoria

Volume 42

Endangered Languages and Languages in Danger. Issues of documentation, policy, and language rights
Edited by Luna Filipović and Martin Pütz

Endangered Languages and Languages in Danger

Issues of documentation, policy, and language rights

Edited by

Luna Filipović
University of East Anglia

Martin Pütz
University of Koblenz-Landau

John Benjamins Publishing Company
Amsterdam / Philadelphia

Table of contents

Foreword

There is general consensus among linguists and language experts that slightly more than 7,000 languages (Ethnologue 2015)[1] are spoken by 7 billion people around the world today and that half of them are under threat of extinction within fifty to one hundred years, a dramatic change in human history. Today at least 3,000 mainly indigenous or local languages are endangered, seriously endangered or dying in many parts of the world. Our planet seems to exhibit an astonishing ratio of speakers to their languages: 97% of the world's people speak about 4% of the world's languages (UNESCO 2003).

This information informs the premise of the current volume because it focusses our attention on two facets of interest here: endangered languages and languages in danger. We believe it is important to make this distinction and address both themes in a holistic approach. *Endangered languages* are those that are moving towards extinction, for a variety of reasons that our contributors discuss, mainly related to diminishing sizes of speaker populations, lifestyle changes and other socio-economic and political factors. *Languages in danger*, on the other hand, refers to the circumstances that create a disadvantaged position for speakers of certain languages when they find themselves within another linguistic environment that speaks a different, majority language. Some of our contributions point out the issues that need to be raised in this context since, technically, many languages, or more precisely their speakers, can find themselves in danger, which can then lead to inequality and injustice. Interestingly, we show that the size of the speaker population does not matter when it comes to languages in danger. Specifically, even speakers of a very populous language can find themselves in danger because of their language within countries where their language is not spoken, thus requiring translation-mediated communication. This status creates a number of barriers for such speakers and may result in serious negative consequences for those speakers. The common thread of argumentation in this volume is that we need to study all the roles that languages and their respective statuses in different contexts play when it comes to social interactions in our multilingual world and the need for peaceful co-existence amidst linguistic and cultural

---

1. The most recent web edition of the *Ethnologue* (2015, 18th edn.) contains information on 7,102 known living languages.

diversity. We hope that the current volume sets out directions for current and future explorations in this vein. The volume examines the causes behind this dramatic loss of linguistic diversity, why this is an issue, how processes of language shift are triggered and what can be done and achieved to document and support endangered, moribund and small languages especially in the context of an ever increasing globalized world. Some of the questions posed in the present volume are, amongst others: How can a minority/indigenous language be maintained in this era of globalization, what are the main reasons for language shift, what do we lose when languages die and what is the role of language policy and planning strategies in multilingual contexts? And finally, what are the benefits of documenting and archiving endangered languages for linguistics, related disciplines and our human cultural heritage in general, especially in the light of new advances in technology and methods of data collection? In this vein, the context of language shift, language threat and loss in multilingual situations are explored, with all the challenges and consequences involved. These are discussed from a variety of perspectives: sociolinguistics, anthropology and the sociology of language including language contact, language ecology, language policy/planning, language rights, and language documentation.

The collection of contributions included in this volume was originally presented at the *36th International LAUD Symposium* on Endangered Languages, which took place on March 31 – April 3, 2014 at the University of Koblenz-Landau (Landau campus). The chapters included for publication here are a small selection from those presented at the conference. A second collection of papers stemming from the same conference was edited by Martin Pütz and Neele Mundt (2016) and is entitled "Vanishing Languages in Context: Ideological, Attitudinal and Social Identity Perspectives" (Peter Lang).

We are indebted to many who have helped us in compiling this volume and whom we hereby acknowledge. First of all, thanks are due to the many internationally well-known conference contributors for their stimulating discussions in a very pleasant atmosphere and to the authors in the present volume, who have responded with professionalism to all the requests that have been made of them. Furthermore, we would like to express our gratitude to the external reviewers who dedicated their time and expertise to reviewing the papers and who gave constructive feedback to the authors. We thank the Series editors, Ana Deumert and Kristine Horner, for the very useful comments on all the chapters. Our gratitude also goes to the excellent John Benjamins production team, in particular Kees Vaes and Patricia Leplae, who never kept us waiting for an answer to a question or a solution to a problem.

Luna Filipović
Norwich, UK

Martin Pütz
Landau, Germany

Introduction

Endangered languages and languages in danger

Luna Filipović and Martin Pütz
University of East Anglia / University of Koblenz-Landau

1. Preliminary remarks

This collective volume brings together the latest research on language endangerment and language rights. It creates a vibrant, interdisciplinary platform for the discussion of the most pertinent and urgent topics central to vitality and equality of languages in today's globalized world. The novelty of the volume lies in the multifaceted view on the *variety of dangers* that languages face today, such as extinction through dwindling speaker populations and lack of adequate preservation policies or inequality in different social contexts (e.g. access to justice, education and research resources). There are examples of both loss and survival, and discussion of multiple factors that condition these two different outcomes. We pose and answer difficult questions such as whether forced interventions in preventing loss are always warranted or indeed viable. The emerging shared perspective is that of hope to inspire action towards improving the position of different languages and their speakers through research of this kind.

This Introduction is not meant as a general survey of the field of the Endangerment of Languages; for that, readers are referred to some of the most recent monographs and volumes that appeared on the topic and related issues, e.g. Austin & Sallabank 2011b; Crystal 2000; Evans 2010; Gippert, Himmelmann & Mosel 2006; Grenoble & Whaley 1998; Grenoble & Whaley 2006; Harrison 2007; Jones 2015; Mihas et al. 2013; Nettle & Romaine 2000; Thomason 2015; Putz & Mundt 2016, etc.).[1] Rather, we offer an interdisciplinary discussion of topics that are related to the key questions in the field and that our contributors address from a whole array of different perspectives, linguistic, socio-cultural and

---

[1] The *Cambridge Handbook of Endangered Languages* edited by Peter Austin and Julia Sallabank (2011b) is certainly the most comprehensive and up-to date collection comprising 23 chapters covering all major areas pertaining to the topic of Endangered Languages.

psychological. This includes some major core thematic areas such as language endangerment, language ecology, language policy, ethnolinguistic vitality, linguistic human rights, language shift and language documentation.

2.   Endangerment of languages and language ecology

The starting point for the study of "endangered languages" can be traced back to Hale's "call to arms" (Simons & Lewis 2013: 17) in 1992 when the loss of linguistic diversity and the areas of language revitalization and documentation aroused the interest of linguists and fieldworkers alike.[2] Hale at the time (1992) described language loss as part of a much larger process of the decay of cultural diversity in which politically superior languages and cultures simply endanger indigenous and local languages and cultures, placing them in an aggravated situation.

How do we define an "endangered language" and how can we tell when a language is endangered? A number of useful suggestions have been put forward to describe the concept from a number of different viewpoints. One brief, but useful definition addressing issues such as numbers of speakers, linguistic usage[3] and intergenerational transmission is provided by UNESCO (2003):

> A language is endangered when its speakers cease to use it, use it in fewer and fewer domains, use fewer of its registers and speaking styles, and/or stop passing it on to the next generation.

A similar, more recent account is given by Thomason (2015: 4) who focusses on language as not being learned as a first or second language anymore, thereby becoming a *moribund* language:

> A language is clearly endangered when it is at risk of vanishing within a generation or two – that is when its last fluent speakers are elderly, when few or no children are learning it as a first language, and when no one is learning it as a second language. Some experts call a language MORIBUND when it is no longer being learned as a first language: a language that is not being transmitted to younger generations cannot outlive the last generations of native speakers.

---

2.   "The world's languages in crisis" (Krauss 1992) was the main topic discussed at a Symposium on Endangered Languages at the annual meeting of the Linguistic Society of America (Simons & Lewis 2013: 3).

3.   Some patterns characteristic of endangered and dying languages refer to a (i) restriction of the vocabulary, (ii) simplification or generalization of inflections, and the (iii) movement of phrases from one position to another (Mesthrie et al. 2009: 256).

The most reliable indicators of language endangerment which seem to be inherent in most of the volume chapters thus include a conglomerate of interrelated factors, i.e. (i) the overall number of speakers (especially speaker age), (ii) intergenerational transmission, (iii) loss of L1 in linguistic and social usage (domains) and (iv) globalization which may all lead to language shift, loss and finally to the "death" or extinction of a language.

These complex relationships between speakers and their languages in their social, cultural, economic and political contexts can be captured by the concept of *language ecology* which is a "dynamic system consisting of a number of inhabitants and meaningful interconnections between them" (Mühlhäusler 1997:5).[4] Language is not isolated from other social, cultural and ecological factors in which it is embedded, but it interacts with them in complex ways. We shall see later that the term "language ecology" in its multifaceted dimension pertains to all contributions in the volume. For example, of particular interest to Gómez' study (this volume) is the observation that among the Yanomae-speakers of Brazilian Amazonia the close relationship between language, culture and the environment (habitat), i.e. their language ecology, is revealed in the use of metaphor, metonymy, and euphemism in quite a number of linguistic expressions such as words and phrases. Likewise Ellis' paper (this volume) strengthens the position that diverse sociolinguistic ecologies emerging from the socio-historical and cultural evolution of languages and dialects in contact (e.g. Saipan Carolinian in Micronesia) have greatly enlarged our understanding of how human language works and how it is conceptually organized.

From the perspective of language ecology, languages somewhat analogous to biological species have a home or natural habitat which is constituted by their speech community.[5] They are dynamic, constantly changing, interacting with and coming into contact with other languages and they exist in a complex social and ecological matrix. They can exist "in a healthy or a degraded habitat, and their transmission from one generation to the next can be threatened, reduced or even fully interrupted" (Harrison 2010:89).

---

4. The metaphor of "ecology" illuminates a number of aspects such as (i) the diversity of inhabitants of an ecology, (ii) the factors that sustain diversity, (iii) the housekeeping that is needed, (iv) the functional interrelationships between the inhabitants of an ecology (Fill & Mühlhäusler 2001:3).

5. Linguists agree that there are correlations between linguistic, cultural and biological diversity. In Papua New Guinea, for example, there is a high number of different biological species and an enormous linguistic diversity including up to 800 languages. The reasons for decline, it is suggested are nevertheless likely to be different (Sutherland 2003).

Mühlhäusler (2003) has proposed the term *ecolinguistics* to account not for the competition between languages, but rather to the interconnections between languages and their environments, i.e. their speakers and the world in which the speakers move. Therefore, from an ecological perspective it is not the size or the number of languages but the meaningful relationships between them and their users' culture that seems most revealing (Grenoble 2011: 31).

Mufwene (2001), however, who also advocates an ecological approach to language evolution, criticizes "self-proclaimed ecolinguists" (Mufwene, this volume) who have ignored the fact that language(s) emerged foremost as a communication device intended to help humans adapt more successfully as groups to ever changing physical and social ecologies. He uses an analogy with tools that are rusting to depict the process of language attrition (Mufwene, this volume):

> …, knowledge of them may become rusty, when members of the relevant populations have not used them for a while, which is what language atrophy or attrition is. The experience is indeed comparable to a metallic tool becoming rusty and potentially less useful.

We may conclude here that an ecological approach to language endangerment, shift and loss of languages is certainly useful and well-placed within the theme, but that it has been evaluated differently depending on whether planning for diversity and a focus on functional links between languages is of primary concern (Mühlhäusler), or whether there is an emphasis of language ecology in the context of the evolutionary aspects of human cultures in dealing with language shift and loss (Mufwene). In a way, the apparently two contradictory approaches can also be observe in the wider context of language conflict seen from the viewpoint of the ideology-based rationalist and romantic models (Pütz 2007). Whereas the rationalist model views language as a medium or tool of rich communication, the romantic model stresses the identification of language and culture and consequently sees language as a functional means of local self-expression and self-identification.

3. Why are endangered languages worth saving (Or what is lost when languages die)?

Researchers documenting and studying language endangerment in a variety of contexts (e.g. Austin & Sallabank 2011b; Harrison 2010; Thomason 2015) seem to agree that there are a number of cogent reasons to maintain, safeguard and document vanishing languages. So what do we lose when languages die? Thomason (2015) poses the question of what the *community* loses and what *science* may lose

with reference to the natural world, e.g. folk medicines and traditional foods used by speakers of endangered languages.

One of the most pertinent reasons as to *why* languages should be saved and maintained is the assumption that language and culture are intimately connected to each other so that (most) members of endangered-language communities and also (most) linguists believe that their cultures cannot fully survive if their languages are moribund or become extinct. The loss of languages at the same time diminishes the cultural diversity of our planet since language is the major transmission tool for culture. No doubt, language is the primary means whereby people conduct their social lives. When it is used in contexts of communication, it is bound up with culture in multiple and complex ways. Accordingly, the argument goes, language expresses cultural reality because people share knowledge about the world, attitudes and belief systems and they also create experience through language. In other words, language embodies and symbolizes cultural reality (Kramsch 1998). As Evans states "language diversity … is intimately tied up with the great plasticity of human experience" (2010: 155) suggesting a co-evolution of language, culture and thought. As communities lose their language they often also lose parts of their cultural traditions or cultural heritage expressed through language in the form of oral history, epic tales, songs, narratives, myths, wordplay and poetry that are not easily translated into other languages especially as the vast majority of human languages have never been written down, let alone documented. These verbal arts exist only in memory and are especially vulnerable to forgetting as languages face the danger of extinction.

In addition to asking what *community* loses in terms of cultural resources, it is therefore appropriate to also ask what *science* may lose. In other words, language loss as a threat to our understanding of human history, human cognition, and the valuable practical knowledge of the natural world (Thomason 2015: 94ff.), since our access to these spheres of knowledge are mainly via language. Linguistic diversity is strongly linked to historical connections among human populations, as well as evidence for ancient population movements. As a case in point, Thomason (2015: 96) refers to Native American Algonquian languages which would have been quite unknown to us in terms of their numbers and geographical spread. Had historical linguists not been able to compile and compare lists of words, phrases and grammatical features it would not have been possible to gain knowledge about the languages' typology and the historical connections between them. History also plays a decisive role in Hein van der Voort's account of Aikanã, an endangered language in Rondônia, Brazil (this volume). He shows that ongoing efforts to document and describe the language as well as the culture and ethno-historical context of the Aikanã can be adapted to the community's needs and do also play a role in maintenance and preservation.

Furthermore, it is a well-known fact that languages also uncover some constraints as well as some possibilities of human cognition and the architecture of human thought (Harrison 2007: 18). Namely, if we manage to identify universal properties of the vastness of all human languages, we may be able to learn about some crucial properties and mechanisms of the human mind (see Hawkins 2014 on cross-linguistic variation and universals in language processing). One thing that is certainly needed is, as Harrison (2007: 19) states, "the oddest, quirkiest, and most unusual languages and words to test our theoretical models". For example, without having studied and documented the endangered language Urarina (Peru), the linguistic world would not have come to know that an O-V-S word order exists in a language (Harrison 2010: 19) and that we can explain why such word orders are possible but comparatively less frequent in the world's languages (see Hawkins 2014 for an explanation). Similarly, we are informed that Ubykh, an extinct Caucasian language has an unusually large repertoire of eighty-four distinct consonants (phonemes) and only two phonologically distinct vowels (Austin & Sallabank 2011a: 6), which makes our knowledge about the variation in the phonetic repertoire of the world's languages more complete. In other words, instead of speculating what linguistic phenomena would be possible or impossible (as some linguistic theories in the past have done) we have evidence of what occurs in human languages with higher or lower frequency; in other words, we are able to form our knowledge base in a more reliable and accurate fashion.

In addition to cultural heritage, historical connections and human cognition it is also our scientific knowledge about humanity and the natural world that is affected when vulnerable languages become extinct. By 2100, more than half of the more than 7,000 languages spoken on Earth – many of them not yet recorded – may disappear, taking with them a wealth of knowledge about history, culture, the natural environment, and the human brain (Harrison 2007: 15).

4.   Causes of language endangerment and disappearance: Language shift

The causes of language shift and consequently language endangerment and loss are diverse and complex. A vulnerable language disappears and eventually becomes extinct when for various reasons its speakers disappear (due to massacres, starvation, diseases) or when they shift to speaking another language by neglecting the use of their heritage language/dialect. Historically, in colonies, and elsewhere where speakers of different languages have come into contact, some languages have been considered superior to others and often one language would attain a dominant position in a community. In most cases social-cultural, political and economic pressures may lead speakers to give up their minority tongue in favor of

a dominant majority language used by an economically and politically more powerful group. Speakers of endangered minority languages may themselves come to associate their language with negative values such as poverty, illiteracy and social stigma. Consequently they wish to adopt the dominant more powerful language which is associated with social and economical progress and modernity.

Thomason (2015: 18ff.) assesses a number of defining factors that contribute to language endangerment such as genocide and fatal diseases, conquest, economic pressures, a melting-pot ideology, language politics, and negative or indifferent attitudes towards a community's vulnerable language. Today, increased migration and rapid urbanization often go hand in hand with the loss of traditional lifestyles or local language ecologies and there is a strong pressure to make use of a dominant or powerful language to engage in socio-economic advancement. So one of the main reasons why language shift occurs leading to language loss is when speakers refrain from using their own heritage language in favor of the language used by the politically and ideologically preferred language or dialect. Even if studies on the effect of globalization on endangered languages are quite scarce (cf. Austin & Sallabank 2011a), there is general agreement that language shift and potential loss is frequently caused by socioeconomic factors: a shift to one of the European languages in a globalized context, i.e. English or Spanish, may provide speakers with better chances of employment, access to international networks and benefits of intercultural communication. However, it is not only an issue of the modern world today that globalization and development have an impact on language shift and language loss. With reference to the Asian context, Lim (this volume) demonstrates that long before today's globalization, the English language was already a key player in language shift scenarios, e.g. during colonial rule, when for example the Peranakan community (Malaysia), with pro-British tendencies and access to English-medium communication, shifted quite early and swiftly to English.

Some researchers have noted that it is not only the foremost official European languages such as English, French or Spanish which impose a language threat to local languages, but rather frequently also lingua francas or Languages of Wider Communication (LWC) such as Lingala, Wolof or Swahili in Africa. In Tanzania, for example, it is the African language Swahili and not the global language English which currently represents a major threat to linguistic diversity in Tanzania (Rosendal, this volume). Similarly, in Kenya there is a rural shift from indigenous languages to Swahili even in the home domain which is facilitated by its widespread use in urban contexts (Gibson & Bagamba, this volume). The authors suggest this happens because Swahili is a language which is not perceived as belonging to one particular ethnic group, and therefore not automatically indicating identity shift. An interesting ecological situation describing community efforts in

an attempt to avoid language shift refers to the case reported by Heine et al. (this volume) where language replacement by a lingua franca does not take place in the ethnic community of the Akie in Tanzania. The Akie represent a "defensive" culture in a sense that "they themselves are setting boundaries vis-a-vis their human environment, trying to keep their culture 'clean' and to prevent outsiders, and especially always the Maasai, from having access to it". Therefore, one effect of this behavior is resistance to language shift and the preservation, maintenance and defense of an indigenous language (i.e. Akie).

5.   Language policy and linguistic human rights

According to Romaine (2002:1) fewer than 4 per cent of the world's languages have any kind of official legal status in the countries where they are used as media of communication. The fact that most languages have not been documented, officially not implemented in the constitution of a nation, limited to home functions in the community, and spoken by very small fractions of people reflects the balance of power in the global linguistic market place.

Over the last few decades, language policy (and planning) has developed into a major discipline, drawing on research and practice in many multilingual and multicultural nations around the word (Spolsky 2012).[6] Sallabank (2011:278) draws a relevant conceptual difference in this context between policy and planning in a sense that policy is geared towards top-down decision-making processes, while planning is usually used with bottom-up or grassroots measures to maintain and support languages; the traditional frameworks of language planning include corpus planning, status planning, language-in-education planning and prestige planning the conceptual characteristics of which, however, will not be elaborated on at present.

Relevant to policy-making is also the study of language attitudes. Opinions, ideas and prejudices that speakers have with respect to languages and cultures in everyday life, education, politics and science play an important role in shaping the dynamics and influencing the fate of a language and hence this represents one of the central issues in language policy and planning strategies. As Thomason (2015:27) points out in regard to attitudes pertaining to indigenous minority language speakers: "even without direct pressure from a dominant speech community, the minority community itself comes to view its own language with

---

6.   The *Cambridge Handbook of Language Policy* (2012) edited by Bernard Spolsky (one of the contributors to this volume) is a complete "state-of-the-field" survey, covering, inter alia, language practices, attitudes and activities associated with indigenous and endangered languages.

indifference or contempt", thereby diminishing the chances for long-term survival. This underlines the importance of prestige planning aimed at promoting a positive view of a language, which can be crucial for the success of language revitalization measures. Most often, language policy and planning failure in most parts of the world are due to insensitivity to the potential ramifications of the language planning context that lacks an informed holistic approach. Very often, for instance, policies are made that take certain major languages into account, but at the same time completely ignore indigenous and endangered languages. Some of these kinds of scenario have resulted in controversies with regard to linguicism and linguistic genocide (Phillipson, Rannut & Skutnabb-Kangas 1995: 9) and the degradation and low regard of a person's identity and ethnic identification.

Identity and ethnicity are also strongly related to a person's right to use his or her language, e.g. in legal, judicial and administrative acts, language education and the media both at an individual and collective level. In this vein, language policy is also intimately tied to the right to speak one's own language. The fact that minority rights should also be treated as human rights has been on the agenda of the language endangerment movement since the publication of Skutnabb-Kangas and Phillipson's (1995) volume on "Linguistic Human Rights" sub-titled "Overcoming linguistic discrimination". Linguistic rights at an individual level mean that (1995: 2)

> everyone can identify positively with their mother tongue, and have that identification respected by others, irrespective of whether their mother tongue is a minority language or a majority language. It means the right to learn the mother tongue, including at least basic education through the medium of the mother tongue, and the right to use it in many of the (official) contexts…Restrictions on these rights may be considered an infringement of fundamental LHRs.
>
> (Linguistic Human Rights)

From the perspective of linguistic human rights, Phillipson, Rannut and Skutnabb-Kangas (1995) argue that individuals and groups are often treated unjustly and are suppressed by language or because of language. They are thus prevented from enjoying other human rights, "including fair political representation, a fair trial, access to education, access to information and freedom of speech, and maintenance of their cultural heritage" (1995: 2). For example, Austin and Sallabank (2011a: 9) cite the example of Kurdish where the language for a long time was denied existence and use in the state of Turkey. The language was banned until 1991, and today it is barred in schools, parliament and other official settings on the grounds that it would divide the country along ethnic lines. Although the situation has improved in recent years, at present, the use of languages other than Turkish in education, politics and the broadcast media is – with minor exceptions – still prohibited.

People who are not fluent in national or official languages need access to services such as education, the media and the justice system, and inadequate translation might deny them access to justice. As we indicated in the Preliminary Remarks section of this Introduction, any language can be in danger of being discriminated against. This directly impacts the status and positions of its speakers in various circumstances, such as access to legal, educational and social services. For instance, Davidson (2000) has shown that interpreter-mediated doctor-patient interactions create significant disadvantages for the patients who have to communicate through interpreters who also act as institutional gatekeepers. It is more often the interpreter that answers the patients' questions rather than the doctor and a great many of the patient's questions remain unanswered under these circumstances. It seems obvious and important that we insist on improvements in interpreter training as well as raise awareness among the medical personnel (and that of other public services) about these issues, especially in the light of our ever increasing multilingual working and social environments.

It is also of relevance to reiterate that it is not necessarily the small and diminishing number of speakers that puts a language in danger. A language with great many speakers can be in danger because the surrounding circumstances lead to a disadvantageous status of its speakers. The reason for this phenomenon is the fact that communication through an interpreter where one's language is not understood directly by the majority of the participants in a communicative situation (e.g. the judge, the jury, the lawyers, the witnesses in an English-speaking courtroom) always carries a certain disadvantage for the speaker of that language. Research has shown that even so-called "big" languages, i.e. very widely spoken by vast populations around the world (e.g. Spanish) can represent a significant disadvantage in an English-speaking legal environment (see e.g. Berk-Seligson 2002; see also Gibbons 2003 for a general discussion and examples on language and disadvantage). However, it is often the case that the endangered languages are also the ones that are most in danger, as Eades (2000, 2002) has shown in her study of the disadvantaged status of Aboriginal languages and the ensuing consequences in Australian courts (see also Berk-Selikson 2008 on language and access to justice by Quichua of Ecuador). In this vein, Hales and Filipović (this volume) discuss communication problems that arise in a multilingual legal context (police interviews and the legal process in general, including both the US and the UK context) such as the need for language support and translation in criminal and immigration proceedings. With reference to real life interviews Hales and Filipović try to assess the difficulties that non-English speakers face in an English-speaking justice system; they argue that people's language rights can indeed be endangered and their treatment seen as lacking in equality, and additionally

contributing to people feeling extreme anxiety and tension of being in danger due to not understanding and not being understood.

Finally, we would like to emphasize the relevance of language policy and planning with reference to language rights when it comes to Sign Languages, i.e. communication in non-oral modalities. Jones (this volume) refers to the extremely low number of new learners of British Sign Language (BSL) and the scarce language planning efforts towards safeguarding the language revealing disregard of deaf people's human rights. Jones even talks about language inequality and linguistic imperialism (Phillipson 1992) which stem from "ignorance of modern research" and surdism (in which deaf people are normalized to be as hearing as possible).

6. Assessing ethnolinguistic vitality status

When we want to assess a language's status or determine whether or not a language is endangered it is useful to make reference to the factor of ethnolinguistic vitality. A first account of vitality has been given by Giles et al. (1977) who define the term as "that which makes a group likely to behave as a distinctive and collective entity within the intergroup setting" (1977: 308).[7] The more vitality an ethnolinguistic group enjoys, the more it will be able to use its own language so as to survive and thrive as a collective entity. According to Giles et al., three dimensions of socio-structural and functional variables influence the vitality of ethnolinguistic groups: (i) demographic, (ii) institutional support, and (iii) status. Furthermore, Fishman (1972) argued that especially interaction networks that employ the language for one or more vital language functions is an important indicator of ethnolinguistic vitality. The more speakers of a language variety and the higher their status, the greater the group's language vitality and the greater the chance for linguistic survival. Due to the ever increasing factors affecting language vitality of an endangered language, the analysis of more complex categories have proved to be necessary. For this purpose linguists have suggested different scales of endangerment ranging from the state of an *endangered,* vulnerable language all the way to a *not endangered,* safe language as the two endpoints of a continuum. Language endangerment, therefore, is a matter of degree. At one end of the scale are languages that are vigorous and safe, and perhaps are even increasing in numbers of speakers or domains of usage, but nevertheless exist under the shadow of

---

7. In 2011, an entire issue of the *Journal of Multilingual and Multicultural Development* was devoted to the study of ethnolinguistic vitality (Vol. 32(2)), with several authors presenting their own tools for measuring language vitality.

a more dominant language. At the other end are languages that are on the verge of extinction (that is, loss of all individuals who continue to use and identify the language as being related to their identity). In between there are many differing degrees of greater or lesser vitality.

In order to take precise account of the ethnolinguistic vitality of a language, a fine-grained classification of language endangerment is provided by UNESCO's "Language Vitality and Endangerment Framework" (2003), which lists six categories of languages according to their status: *safe*, *vulnerable*, *definitely endangered*, *severely endangered*, *critically endangered*, and *extinct*. Similarly, Grenoble and Whaley (2006: 18) propose a comparable six-way distinction to capture different stages of endangerment, i.e. *safe, at risk, disappearing, moribund, nearly extinct* and *extinct*. In line with these descriptive labels illustrating the status of endangerment, linguists have assessed the accuracy of these criteria by applying a multitude of ethnolinguistic vitality frameworks to endangered languages, e.g. Fishman's (1991) "Graded Intergenerational Disruption Scale (GIDS) which served as the best-known evaluative framework for nearly two decades, as well as Lewis and Simon's (2010) extended version of GIDS (EGIDS) amongst others.

As mentioned above, ethnolinguistic vitality is one of the key concepts when it comes to assessing the vitality and strength of an endangered language. It is generally agreed by linguists that roughly the following nine criteria may determine the vitality of a language, such as changes in the number of speakers or in the use of the language in certain domains or functions (UNESCO 2003):

–   Intergenerational language transmission
–   Absolute number of speakers
–   Proportion of speakers within the total population
–   Shifts in domains of language use
–   Response to new domains and media
–   Availability of materials for language education and literacy
–   Governmental and institutional language attitudes and policies including official status and use
–   Community members' attitudes toward their own language
–   Amount and quality of documentation

Certainly, the nature of the speaker base (numbers and proportions of speakers); the domains of language use where the native language was once secure (churches, schools, the cultural sphere and, most important, the home) and the institutional support and attitudes can be singled out as the main instigators of language endangerment/vitality (Grenoble 2011: 38). The key factor in assessing the relative safety of an endangered language is the degree to which intergenerational

language transmission of the language remains intact. For instance, if children do not acquire an indigenous language as first medium of communication anymore, then there will be a tendency towards lesser usage of the language and concomitantly the danger of that language becoming extinct in the near future.

7.   Responses to language endangerment

Linguists, language activists, members of endangered language communities, governments, non-governmental organizations, and international organizations such as UNESCO and the European Union are actively working to save and stabilize endangered languages. Once a language is classified as endangered, there is a number of 'responses' that can be taken in order to stabilize or rescue a vulnerable language (see contributions on 'Responses to Language Endangerment' in Mihas et al. 2013). According to Grenoble (2011:43) responses can be twofold: documenting endangered languages while still possible, and restoring and reviving them.

7.1   Language documentation

Creating lasting documentations is seen as one major response to the challenge of the dramatically increased level of language endangerment although it should be said from the outset that language documentation is in principle quite general, not specific to endangered languages. One of the major studies focusing on the "Essentials of Language Documentation" is Gippert, Himmelmann and Mosel's (2006) volume, which discusses a number of issues relevant to the topic such as data collection and processing, fieldwork ethics, ethnography, linguistic annotation, orthography, archiving and mobilizing language documentation with multimedia, etc. Himmelmann (2006), one of the main proponents of documentary linguistics research, defines language documentation as "a field of linguistic inquiry and practice in its own right which is primarily concerned with the compilation and preservation of linguistic primary data and interfaces between primary data and various types of analyses based on these data" (2006:1). More precisely and brief, Woodbury (2011:159) speaks of language documentation as "the creation, annotation, preservation and dissemination of transparent records of a language". In line with the definition, Thomason (2015:113) refers to *creating the records* essentially as collecting field data and a language corpus; *annotation* involves setting up a list of metadata indicating the date, place, speakers, genre, context, etc. *Preservation* means depositing the data in digital archiving systems,

and *dissemination* means making the annotated records available to linguists and ethnographers, and other people of the endangered-languages movement. What makes the record *transparent* is annotation of the material. From a historical perspective, Austin (this volume) discusses the emergence and development of language documentation during the past 20 years and observes a recent shift towards recognition of diversity (of projects, goals, participants and outcomes), multi-code contexts and the changing role of technology. An interesting example of endangered language documentation dealing with lexical erosion in Palikur, an Amazonian language, is provided by Nemo and Cristinoi (this volume). They focus on the necessity to systematically document those parts of the lexicon which are the most likely to disappear, i.e. animal plants, ritual practices, mythical entities, traditional medication, etc.

On this occasion we do not delve into the details with regard to the technical implications of language documentation.[8] However, one issue which should be a major concern of linguists and language planners is the status and role of the community members. As Grinevald and Bert (2011:45) suggest "it is (therefore) worth reminding ourselves that speakers are indeed the source, not to say the heart and soul of it all". The ideal situation is one in which the fieldworkers are themselves members of the community, trained in linguistics (likewise in ethnography). A linguistic fieldworker in turn will train one or more community members to participate actively in the research, perhaps by interviewing elders or by recording and translating stories and songs. The relationship between the researcher or fieldworker and the community member is often seen from an ideological perspective in a sense that power relations between the researcher and the researched are identified. From this perspective, Grinevald and Sinha (this volume) speak of "North-South relations in linguistic science" questioning the motivations and practices of endangered language documentation research (ELDR):

> When endangered languages are viewed primarily through the lens of heritage, and the everyday language practices of the communities that speak (or used to speak) them are neglected, these communities themselves are de-privileged in the discourse of Endangered Language Documentation Research (henceforth ELDR). In this case, we can speak of a "heritage ideology" which is reinforced by influential (and often uncritically accepted) conceptions of linguistic science and scientific method.

---

8.   Here the reader is referred to the five chapters on „language documentation" (Part II) in the *Cambridge Handbook of Endangered Languages* (Austin & Sallabank 2011b) as well as to the volume "Essentials of Language Documentation" (Gippert, Himmelmann & Mosel 2006) which is an introduction to basic practical and theoretical issues in language documentation.

Generally, they criticize the view that languages are primarily "science heritage" resources, and only secondarily vital to the benefit of community members and of the survival of indigenous communities. Contrary to a "heritage ideology", a rather positive example of a favourable attitude towards the strengthening of indigenous languages and cultures is provided by Cabral, Sampaio and da Silva Sinha (this volume). They report on recent governmental policy efforts in Brazil which have contributed positively to indigenous educational programs and documentation and which also aim at training indigenous people as teachers and researchers.

7.2   Language revitalization

Language documentation is closely linked to the field of language revitalization, also referred to as language revival or reversing language shift (RLS). Revitalization is applied to "the phenomenon of attempting to bring endangered languages back to some level of use within their communities (and elsewhere) after a period of reduction in usage" (Hinton 2011: 291). Furthermore, it is the process by which a language community through political, community, and educational means attempts to increase the number of active speakers of the endangered language. It has been pointed out (for example by Spolsky, this volume) that there has only been one successful instance of a complete language revival process, that of the Hebrew language, now the national language of Israel, creating a new generation of native speakers without any pre-existing native speakers as a model.

Revitalization efforts are strongly connected with intergenerational transmission categorization scales. The most influential of such scales is certainly Fishman's (1991) eight-point – *Graded Intergenerational Disruption Scale* (GIDS) which focuses on the key role of intergenerational transmission in the maintenance and preservation of a language.[9] It also describes and explains stages in reversing language shift when linguists and language activists are trying to turn threatened languages into safe ones (Simons & Lewis 2013: 5). If children do not learn a language from their parents, it can hardly be guaranteed that they in turn will be able to pass the language on to the second generation of speakers, i.e. their children. Fishman's scheme at the same time indicates specific activities and measures for language shift to be reversed at each of the 8 points of the GIDS scale.

---

9.   Another framework of Ethnolinguistic Vitality by Landweer (2012), i.e. *Indicators of Ethnolinguistic Vitality* (IEV) is applied by Smith (this volume) in her evaluation of the endangered language Papana (New Guinea).

Twenty years later, Lewis and Simons (2010) critiqued the GIDS on the grounds that it was not comprehensive enough and proposed the multi-dimensional Extended GIDS (EGIDS) which consists of 13 levels with each higher number on the scale representing a greater level of disruption to the intergenerational transmission of the language (Ethnologue) including labels such as "international" (level 1), "educational" (level 4), "threatened" (level 6b), or "dormant" (level 9). Thus, the model is principally based on Fishman's work, but considers an expansion from eight to thirteen levels with the aim to recognize more fine-grained different degrees of vitality over the entire range of the vitality-endangerment continuum (Simons & Lewis 2013: 5). However, as pointed out by Smith (this volume), EGIDS does not necessarily constitute an improvement on GIDS, at least not as far as the Papapana speech community (Papua New Guinea) and social milieu is concerned. Both GIDS and EGIDS consider different factors at different levels within one categorisation scale. To choose a particular level, as the *Ethnologue* does for Papapana for example, would be to ignore other pertinent factors and would thus be misleading, as she argues in this volume. See also Gibson and Bagamba's application of the EGIDS model to endangerment studies in East Africa (this volume).

Interestingly enough, some of the languages being documented by researchers which reveal important insights into how humans communicate in non-oral modalities refer to the status and use of endangered sign languages. Bickford, Lewis and Simons (2015) report on a first attempt to accommodate signed and spoken languages on an equal level using their EGIDS model, i.e. the tool to analyze the status, role and degree of language endangerment. They point out that despite differences between signed and spoken languages, the basic sociolinguistic factors that have an impact on language vitality are substantially similar in the two modalities. A concrete example on how to apply Fishman's GIDS model to British Sign Language (BSL) is provided by Jones (this volume) who claims that according to the factors and dimensions attached to GIDS, BSL can be categorized as *severely endangered*. This is also due to the fact that the number of new learners is extremely low. Furthermore that sign languages are not motherese languages as most deaf children are born to hearing parents who do not sign and therefore language transmission does hardly take place within the family domain. To compensate for this failure, Jones requires language planning initiatives and immediate intervention, against the backdrop of a disregard of deaf people's human rights with "English being the colonial language of oppression".

8.   Critique of the endangered-languages movement

Almost 25 years ago, Krauss (1992) warned the linguistic world to counter-act patterns characteristic of endangered and dying languages, predicting that 50% of the existing languages were doomed to die or in the process of dying. The situation may be even worse, as many as 90 percent of the world's languages could face language extinction by the year 2100; and only 5 to 10 percent were "safe" in the sense of being widely spoken or having official status. According to Simons and Lewis (2013:15), even if on a global scale the threat does not yet reach the pessimistic level suggested by Krauss, the situation as such is still alarming and linguists as well as language activists and planners worldwide have taken a strong stance in trying to halt the looming crisis of language decay and loss. Due to such massive appearance of recent treatments about language endangerment or language death, Duchene and Heller (2007) take some critical distance from this explosion of discursive material and analyze the ways language endangerment ideologies are discursively constructed. In other words, in whose interest is it to mobilize resources around the defense of languages, and why? What are the ideological dimensions of this discourse? What is at stake in these discursive practices and in whose interest is it to promote or contest such discourses.

Despite the fact that most linguists and language activists do indeed strongly support work on endangered languages emphasizing the need to document and revitalize vulnerable tongues, there is also a debate as to why it should be worthwhile interfering with the fate of these languages at all. Not everyone agrees that the worldwide loss of linguistic diversity is something to be counter-acted. For example Mufwene (this volume) rather critically examines the common assumptions about the values of a close relationship between language, culture and identity. He argues that language decay or death is a natural part of the process of human cultural evolution and that languages die for their own reasons. Therefore, he argues, linguists should document languages scientifically, but not interfere with the processes of language shift and loss; thus Mufwene's paper stands in direct opposition to the idea of "language rights".

In a similar vein, from the admittedly "popular press", an article for the American business magazine FORBES entitled "No tears for dead tongues" and authored by the Columbia University linguist McWhorter (2008) advances the argument that the death of a language does not necessarily mean the elimination of a culture and that "keeping a language artificially alive could be worse". And further, in an article entitled "Why save a language?" in the *Sunday Review* of *The New York Times* (2014), McWhorter presents the reader with the following questions: "if indigenous people want to give up their ancestral language to join the

modern world, why should we consider it a tragedy? Languages have always died as time has passed. What's so special about a language?"

Similarly, Romaine (2004) has reviewed a number of articles written by popular commentators and journalists who do not seem to recognize unequal sociopolitical conditions in a society; these are legitimized in the name of "linguistic modernization" and greater "communicative currency" or "languages of wider communication" on the part of the majority-language group (May 2001: 147). Therefore, one major argument of these commentators is rooted in what Romaine calls "the ideology and rhetoric of free market capitalism". She cites Malik (2000), for instance, who claims that the reason why most languages die, is not because they are suppressed, but because native speakers yearn for a better life. Speaking a language such as English, French or Spanish, and discarding traditional habits, can open up new worlds and is often a ticket to modernity. Malik has also argued that it is "irrational" to try to preserve all the world's languages, as language death is natural and in many cases inevitable, even with intervention. In fact, he proposes that language death improves communication by ensuring more people speak the same language. This may, in his opinion, benefit the economy and reduce conflict. Linguists may be trying to preserve the unpreservable, all possessed of a nostalgic view of what constitutes a culture or a 'way of life'. Neither a culture, nor a way of life, nor yet a language, has a 'God-given right to exist'. What if half the world's languages are on the verge of extinction? "Let them die in peace", says Malik (2000). Similarly, the phonetician Ladefoged has "another view of endangered languages" (1992), in agreement with Mufwene's, arguing that language death is a natural part of the process of human cultural development, and that languages die because communities stop speaking them for their own reasons. He argues that linguists should simply document and describe languages scientifically, but not seek to interfere with the processes of language decay and loss, i.e. to be wary of arguments for preserving languages based on political considerations (1992: 809).

So these authors' arguments seem to be based on the view that we should accept changes in language use as a normal process and that extinction is a fact of life. Language evolution is taking place every day; why interfere with it? Language death comes about because people make a free choice to shift to another language. As people are rational beings, who may reasonably be expected to know where their self-interest lies, we, as outside observers, cannot condemn such choices; nor should we intervene in the linguistic market. The decline of some languages is just a side-effect of countless individual choices, and thus is "no more or no less morally significant than a change in the price of fish" (Romaine 2004).

Let us close this Introduction with a somewhat longer quote by Suzanne Romaine (2004) who argues against the views we just presented in this section on

the grounds that these authors almost neglect the workings of political power and ideology in undemocratic societal systems and treatments of human behavior.

> Although some of these critics (*mentioned above*, editors' addition) acknowledge that the rapid decrease in the number of languages over the past few centuries is connected with European colonization of the world and Western economic expansion, they tend to downplay the power imbalances underlying the material, political and economic domination of most of the world's small language communities. This imbalance has allowed a few metropolitan groups a virtual stranglehold upon global resources and global power. Glossing over undeniable disparities in power underlying the history of language shift allows them to ignore the fact that in many cases, language death occurred not because of an increase in the available choices, but because of a decrease in choice brought about by the exercise of undemocratic power. Such power is almost always wielded by denying access to resources from which communities make their living. Languages can only exist where there is a community to speak and transmit them. A community of people can exist only where there is a viable environment for them to live in, and a means of making a living. Where communities cannot thrive, their languages are in danger. When languages lose their speakers, they die.

The editors of this volume firmly believe that the principal ideas, wishes and arguments put forward in the various contributions are offered in good faith, but they need to highlight their belief that it is the right of each individual speaker to have enough freedom of choice: to keep their heritage language and culture, or else to give up their language and shift to another medium of communication (for whichever individual reason this may be) even when language loss and death is likely to occur. Eventually with no doubt, minority-language speakers quite often see their social, cultural and economic advancement in the guise or under the shadow of a politically and economically superior majority language. It is by all means worthwhile to document our linguistic and cultural heritage in all its entirety and leave this information for posterity. As for engineering language survival, the issue is much more thorny and would require stopping or diverting some big societal seismic shifts, where multiple factors contribute to the changes that impact language survival. It will be hard for language specialists to do much if the community that speaks an endangered language is not willing to do so and help should be offered if such efforts are visibly made. One thing that has to be secured though is the assurance and practical implementation of language equality, with respect to languages' political and social status, opportunities and services available and individual attitudes. We believe that volumes such as the current one will contribute to this end by raising awareness about where and why inequality arises, changing biased and discriminating views and motivating further research

and practical actions that benefit both linguistic communities and speaker communities concerned.

References

Austin, P. & Sallabank, J. 2011a. Introduction. In Austin & Sallabank (eds), 1-24.
     doi: 10.1017/CBO9780511975981.001

Austin, P. & Sallabank, J. (eds). 2011b. *The Cambridge Handbook of Endangered Languages*. Cambridge: CUP.   doi: 10.1017/CBO9780511975981

Crystal, D. 2000. *Language Death*. Cambridge: CUP.   doi: 10.1017/CBO9781139106856

Berk-Seligson, S. 2002. *The Bilingual Courtroom: Court Interpreters in the Judicial Process*. Chicago IL: The University of Chicago Press.

Berk-Seligson, S. 2008. Judicial systems in contact: Access to justice and the right to interpreting/translating services among the Quichua of Ecuador. *Interpreting* 10(1): 9–33.
     doi: 10.1075/intp.10.1.03ber

Bickford, A., Lewis, P. & Simons, G. 2015. Rating the vitality of sign languages. *Journal of Multilingual and Multicultural Development* 36(5): 513–527.   doi: 10.1080/01434632.2014.966827

Davidson, B. 2000. The interpreter as institutional gatekeeper: The social-linguistic role of interpreters in Spanish-English medical discourse. *Journal of Sociolinguistics* 4(3): 379–405.
     doi: 10.1111/1467-9481.00121

Duchene, A. & Heller, M. (eds). 2007. *Discourses of Endangerment: Ideology and Interest in the Defence of Languages*. London: Continuum.

Eades, D. 2000. I don't think it's an answer to the question: Silencing aboriginal witnesses in court. *Language in Society* 29(2): 161–195.   doi: 10.1017/S0047404500002013

Eades, D. 2002. Evidence gained in unequivocal terms: Gaining consent of aboriginal young people in court. In *Language in the Legal Process*, J. Cotterill (ed.), 162–179. Houndmills: Palgrave.   doi: 10.1057/9780230522770_10

Ethnologue: Languages of the World, 18th edn, 2015, M. Lewis, G. Paul, F. Simons & C.D. Fennig (eds.). Dallas TX: SIL International. <www.ethnologue.com>

Evans, N. 2010. *Dying Words. Endangered Languages and What They Have To Tell Us*. Malden MA: Wiley-Blackwell.

Fill, A. & Mühlhäusler, P. 2001. Introduction. In *The Ecolinguistics Reader. Language, Ecology and Environment*, A. Fill & P. Mühläusler (eds), 1–9. London: Continuum.

Fishman, J.A. 1972. The relationship between micro- and macro-sociolinguistics in the study of who speaks what langage to whom and when. In *Sociolinguistics*, J.B. Pride & J. Holmes (eds), 15–32. Harmondsworth: Penguin.

Fishman, J.A. 1991. *Reversing Language Shift: Theoretical and Empirical Foundations of Assistance to Threatened Languages*. Clevendon: Multilingual Matters.

Gibbons, J. 2003. *Forensic Linguistics*. Oxford: Blackwell.

Giles, H., Bourhis, R. & Taylor, D. 1977. Towards a theory of language in ethnic group relations. In *Language, Ethnicity and Intergroup Relations* [European Monographs in Social Psychology 13], H. Giles (ed.), 307–348. London: Academic Press.

Gippert, J., Himmelmann, N.P. & Mosel, U. 2006. *Essentials of Language Documentation* [Trends in Linguistics: Studies and Monographs 178]. Berlin: Mouton de Gruyter.
     doi: 10.1515/9783110197730

Grenoble, L. 2011. Language ecology and endangerment. In Austin & Sallabank (eds), 27–44. doi:10.1017/CBO9780511975981.002

Grenoble, L. & Whaley, L. (eds). 1998. *Endangered Languages: Language Loss and Community Response*. Cambridge: CUP. doi:10.1017/CBO9781139166959

Grenoble, L. & Whaley, L. 2006. *Saving Languages: An Introduction to Language Revitalization*. Cambridge: CUP.

Grinevald, C. & Bert, M. 2011. In Austin & Sallabank (eds), 45–65. doi:10.1017/CBO9780511975981.003

Hale, K. 1992. Endangered languages: On endangered languages and the safeguarding of diversity. *Language* 68(1): 1–42.

Harrison, K.D. 2007. *When Languages Die. The Extinction of the World's Languages and the Erosion of Human Knowledge*. Oxford: OUP. doi:10.1093/acprof:oso/9780195181920.001.0001

Harrison, K.D. 2010. *The Last Speakers. The Quest to Save the Most Endangered Languages*. Washington DC: National Geographic.

Hawkins, J.A. 2014. *Crosslinguistic Variation and Efficiency*. Oxford: OUP. doi:10.1093/acprof:oso/9780199664993.001.0001

Himmelmann, N. 2006. Language documentation: What is it and what is it good for? In Gippert, Himmelmann & Mosel (eds), 1–30. doi:10.1515/9783110197730

Hinton, L. 2011. Revitalization of endangered languages. In Austin & Sallabank (eds), 291–311. doi:10.1017/CBO9780511975981.015

Jones, M.C. (ed.). 2015. *Policy and Planning for Endangered Languages*. Cambridge: CUP. doi:10.1017/CBO9781316162880

Kramsch, C. 1998. *Language and Culture*. Oxford: OUP.

Krauss, M. 1992. The world's languages in crisis. *Language* 68(1): 4–10. doi:10.1353/lan.1992.0075

Ladefoged, P. 1992. Another view of endangered languages. *Language* 68(4): 809–811. doi:10.1353/lan.1992.0013

Landweer, M.L. 2012. Methods of language endangerment research: A perspective from Melanesia. *International Journal of the Sociology of Language*. 214: 153–178.

Lewis, M.P. & Simons, G.F. 2010. Assessing endangerment: Expanding Fishman's GIDS. *Revue Roumaine de Linguistique* 2: 103–120.

Malik, K. 2000. Let them die. The preservation of dying languages and cultures is pointless and reactionary. People want to join modernity. *Prospect Magazine*. November 20.

May, S. 2001. *Language and Minority Rights. Ethnicity, Nationalism and the Politics of Language*. Edinburgh: Pearson.

McWhorter, J. 2008. No tears for dead tongues. *FORBES Magazine*, 21 February 2008.

McWhorter, J. 2014. Why save a language? *Sunday Review*. New York Times, 5 December 2014.

Mesthrie, R., Swann, J., Deumert, A. & Leap, W. 2009. *Introducing Sociolinguistics*, 2nd edn. Edinburgh: EUP.

Mihas, E., Perley, B., Rei-Doval, G. & Wheatley, K. (eds). 2013. *Responses to Language Endangerment. In Honor of Mickey Noonan* [Studies in Language Companion Series 142]. Amsterdam: John Benjamins. doi:10.1075/slcs.142

Mühlhäusler, P. 1997. Language ecology – Contact without conflict. In *Language Choices. Conditions, Constraints, and Consequences* [IMPACT Studies in Language and Society 1], M. Pütz (ed.), 3–15. Amsterdam: John Benjamins. doi:10.1075/impact.1.03muh

Mühlhäusler, P. 2003. *Language of Environment. Environment of Language. A Course in Ecolinguistics*. London: Battlebridge.

Mufwene, S. 2001. *The Ecology of Language Evolution*. Cambridge: CUP.
  doi:10.1017/CBO9780511612862

Nettle, D. & Romaine, S. 2000. *Vanishing Voices: The Extinction of the World's Languages*. Oxford: OUP.

Phillipson, Robert. 1992. *Linguistic Imperialism*. Oxford: OUP.

Phillipson, R., Rannut, M. & Skutnabb-Kangas, T. 1995. Introduction. In *Papers in European Language Policy*, T. Skutnabb-Kangas & R. Phillipson (eds), 1–22. Roskilde: Universitetscenter.

Pütz, M. 2007. The dynamics of language policy in Namibia: A view from cognitive sociolinguistics. In *Living through Languages: An African Tribute to René Dirven*, C. van der Walt (ed.), 91–113. Stellenbosch: SUN Press.

Pütz, M. & Mundt, N. (eds). 2016. *Vanishing Languages in Context: Ideological, Attitudinal and Social Identity Perspectives* [Duisburg Papers on Research in Language and Culture]. Frankfurt: Peter Lang.

Romaine, S. 2002. The impact of language policy on endangered languages. *International Journal on Multicultural Societies* 4(2): 1–28.

Romaine, S. 2004. Endangered languages: There's nothing benign about benign neglect. *Language Politics and the Politics of Language* (in Education), 17–19.

Sallabank, J. 2011. Language policy for endangered languages. In Austin & Sallabank (eds), 277–290. doi:10.1017/CBO9780511975981.014

Simons, G.F. & Lewis, P. 2013. The world's languages in crisis. A 20-year update. In Mihas et al. (eds), 3–19. doi:10.1075/slcs.142.01sim

Skutnabb-Kangas, T. & Phillipson, R. (eds). 1995. *Linguistic Human Rights: Overcoming Linguistic Discrimination*. Berlin: Mouton de Gruyter. doi:10.1515/9783110866391

Spolsky, B. (ed.). 2012. *The Cambridge Handbook of Language Policy*. Cambridge: CUP.
  doi:10.1017/CBO9780511979026

Sutherland, W. 2003. Parallel extinction risk and global distribution of languages and species. *Nature* 423: 276–279. doi:10.1038/nature01607

Thomason, S. 2015. *Endangered Languages. An Introduction*. Cambridge: CUP.

UNESCO Ad Hoc Expert Group on Endangered Languages. 2003. *Language Vitality and Endangerment*. Paris: UNESCO.

Woodbury, A.C. 2011. Language documentation. In Austin & Sallabank (eds), 159–186.
  doi:10.1017/CBO9780511975981.009

Perspectives on endangerment

Ideology, language policy and language rights

North-South relations in linguistic science

Collaboration or colonialism?

Colette Grinevald* and Chris Sinha
University of Lyon 2 / Hunan University

In this chapter, we attempt to unmask the ideological bias inherent in influential conceptions of the methods, motivations and practices of Endangered Language Documentation Research (ELDR) by addressing the unequal exchange that frequently characterizes the relationship between the linguistic researcher, on the one hand, and the language community and, in some cases, local researchers, on the other. We highlight the extent to which common answers to the question "Why document endangered languages?" suppress the sociocultural and historical relations within which ELDR practices are situated. We review the historical evolution of the conceptualization of language documentation research, and its relationship to language preservation and revitalization. We ask what it is that makes ELDR scientific, critically analysing the models of "language" and of "science" that are frequently deployed in arguments for its importance, and question the value-neutrality of the notion "scientific community" in this context. We suggest that the conjunction of dominant concepts of "language" and "data", and the relations between "international" and "local" Endangered Language Documentation (ELD) researchers, generates an ideological construction of unequal competence that operates to justify unequal North-South exchange relations. We document this claim of unequal and at times abusive North-South exchange with brief, anonymized case studies. We conclude by noting that, in comparison with other social science disciplines, linguistics seems resistant to reflexive and self-critical analysis of its ideological dimension; and suggesting possible ways of raising awareness and generalizing models of good practice.

Keywords: endangered languages, ideology, language documentation, language preservation, linguistic science, North-South relations, revitalization

---

1.   Introduction: Why document endangered languages?

To most linguists, and many non-linguists perhaps, the answer to this question is self-evident. Both existing, and documented but extinct, languages are vital resources for scientists to study the nature of, and the constraints on, language variation; and thus, indirectly, the human language faculty and the human mind more generally. The languages and language varieties of interest to Endangered Language Documentation researchers will not be, or may not be, there to study *in vivo* in the future, and if they are not documented now they will not be available for future generations of scholars. This is the argument most frequently advanced by linguists to convince governments, inter- and non-governmental agencies and other funding bodies that this research enterprise matters and requires funding. Sometimes, the argument is also put forward that language diversity is a value in itself, analogously with biodiversity, in that languages are the most important bearers of the cultural heritage of the communities that speak them. These two arguments are not, however, always seen as necessarily complementary, and when the latter is advanced it is frequently couched in terms of immaterial heritage whose value (analogously, in this context, with material heritage artworks and artefacts) is frequently conceptualized in terms of larger groups – nation states and their citizens, humanity as a whole – rather than for the still-existing speech communities.

We do not challenge the force or validity of either of these arguments in themselves. But we do question the bypassing of unarticulated critical questions – "whose heritage?" and "heritage or contemporary life resource?" – that the focus on scientific, national and human heritage tends to push to the background. When endangered languages are viewed primarily through the lens of heritage, and the everyday language practices of the communities that speak (or used to speak) them are neglected, these communities themselves are de-privileged in the discourse of Endangered Language Documentation Research (henceforth ELDR). In this case, we can speak of a "heritage ideology" which is reinforced by influential (and often uncritically accepted) conceptions of linguistic science and scientific method. The conception of language that informs and sustains the heritage ideology is that of "language as an object", viewed from outside by an (ideally disinterested) scientific observer who is equipped, by both training and material resources, to analyse the language and disseminate the results of her/his analysis to other members of the scientific community.

Again, we do not challenge these conceptions *in themselves*. ELD researchers are subject to the same professional imperative to communicate research findings to the scientific community as researchers in other fields – even more so, perhaps, in the recent past, when ELDR faced the task of legitimizing itself as a recognized

field of scientific activity. Furthermore, the mainstream traditions of linguistic science have long viewed language as a synchronic system, taking a 3rd person perspective, and the working assumptions of many subfields of linguistics, including most approaches to typological linguistics, reflect this conceptualization. However, both theoretically and in relation to ELDR practices, the language-as-object view is constricting and incomplete. Although it may represent the historically dominant paradigms of linguistics, it does not represent the whole of language science; in particular it does not represent those research enterprises that are primarily interested in languages as communicative tools, and in speech/ language as a social practice.

The "language-as-object" view meshes with the heritage ideology because it views a language as a kind of symbolic package handed down from generation to generation, whose transmission is ideally both whole and accurate, with deviations from these normative ideals (such as those arising from language contact or intergenerational differences in contexts of acquisition and use) being viewed as unfortunate noise in the data. The language-as-object view thus also reinforces the choices of data that are privileged in ELDR, that is, monologic speech genres (such as narratives) produced by older speakers, who are viewed as (and often are) more knowledgeable about both language and cultural heritage (Austin, this volume). Again, we do not dispute the importance of such data, but we do emphasize that this focus is informed by the heritage ideology, in which the data archive is a museum of a past and more "pure" era, rather than a record of contemporary appropriation strategies towards the endangered language employed by speakers with varying degrees of knowledge, occupying diverse positions in the society.

The language-as-object view may serve, then, as an ideological support for the conceptual separation of the endangered language from the community that speaks (or is ceasing to speak) it. Just as Richard Dawkins (1976) argued that organisms are merely vehicles for the replication of genes, so communities may be viewed as merely vehicles for the transmission of a language, language variation and language relationships being viewed from an internalist and autonomous perspective, divorced from both communities and cultures. This is not a caricature, as we can see from the controversy generated by recent defences of the argument (which is itself not new: Sharifian 2015) that grammar is significantly culturally motivated (Everett 2005, 2009; Nevins, Pesetsky & Rodrigues 2009). We acknowledge, of course, that the language-as-object view does not *entail* that language be viewed as autonomous from culture and community. Rather, we are suggesting that it may *reinforce* a background presupposition of such autonomy. The consequence is that not only do ELDR practices privilege certain data choices over others, but also the language itself is viewed as free-floating data, ownership *of* which and access *to* which communities neither have nor should have special

rights over. Endangered language data are, in this view, primarily *scientific herit-age* data, in whose use and preservation the main stakeholders are *the scientific community*.

The countries containing "sites of interest" for ELDR are often in the non-Anglophone global "South". The ELD researchers, and the agencies funding the research, are often from the global "North".[1] The relationship between researcher and researched is, therefore, frequently a *North-South relationship*, with all the socio-political complexities that this carries with it. It is also, of course, the case that linguists who are nationals of countries of the "South" also conduct ELDR, sometimes in conjunction with language revitalization programmes, and some-times in collaboration with researchers from the "North". We address the North-South dimension of researcher-community relations in more detail in Section 3; and North-South ELDR research collaborations in Section 4. Here, we wish to stress that the language-as-object view also operates as an ideological support for unequal power relations (and *unequal exchange*) not only *between* the academe and the communities, but also *within* the academe, because of the way in which it valorizes certain kinds of skills over others. In the academic world of linguistic science, "theory" and "pure research" have traditionally been more highly valued than "applied" research; "data collection" is seen as a purely mechanical, cognitive-ly "light" activity; and local and contextual knowledge gained through personal experience of field work is devalued in comparison with theoretical knowledge. All of this amounts to the devaluation, familiar to many experienced field work-ers, of *knowledge as practice* and *practice as knowledge*.

Knowledge as practice/practice as knowledge can be thought of as encom-passing both the "know-how" that underpins field research, and the reflexive stance that is *theoretically* enjoined upon researchers in all disciplines that employ qualitative research methods. It includes, but is not reducible to, knowledge of how to use technical equipment and software tools. Unfortunately, just such a re-duction, that bypasses the difficult and complex questions of the theory and prac-tice of ethnographic research, intercultural communication, field research ethics, indigenous rights and intellectual property, is all too common in the education and training of ELD researchers. We return to these issues in our concluding

---

1.   The terms "North" and "South", like "Western" and "non-Western", are not strictly geo-graphical in this context. They are metonyms for conceptual complexes that differentiate global populations on geo-political and economic grounds, with a component of differential valoriza-tion. For example, the so-called "Fourth World" indigenous minorities of North America may be considered to be part of the global "South", while the non-indigenous majority of, say, Aus-tralia, may be considered to be part of the global "North". We therefore use scare quote marks for these terms (but not for the phrase North-South relations, which should be understood as denoting relations between the global "North" and the global "South").

reflections. In the next section, we examine in greater depth, on the basis of the previous work of the first author, the development over the past twenty years of changing conceptions of ELDR, of the relationships between documentation, description, archiving and revitalization, and of the differing spheres of ideology within which ELDR exists and which impinge upon its theory and practice.

2.   ELDR: Evolving paradigms

At this juncture we would like to propose a set of constructs to clarify the evolving paradigms of the still developing field of ELDR, which should be understood as having contributed to the ongoing collective debates and publications of the last decade (such as the collections in Brenzinger 2006 and Austin & Sallabank 2010). This set of constructs, which encapsulate changing ELDR paradigms, will serve as a backdrop for our later discussion of North-South relations in endangered language research.[2]

The formula that we propose in order to capture the evolution of the new subdiscipline of ELDR is 'D–D.A+R', in which the linear order – *Description*, then *Documentation* and *Archiving*, then *Revitalization* – reflects the chronological order of elaboration of these subfields; and the symbols used to link them are meant to match familiar linguistic glossing conventions marking different types of relations between elements. The symbol '–' in 'D–D' is meant to resemble a morphemic segmentation, meaning 'description *and* documentation'; '.' in 'D.A' to show the unit formed by 'archiving *of* documentation'; and '+' in 'D.A+R' to mark the addition over time of revitalization as an activity systematically *incorporated in* the complex of practices constituting ELDR. This last link of 'D–D.A plus revitalization' can also be formulated more explicitly as D–D.A *'FOR'* revitalization.

Within modern linguistics, the activity of *description* was long regarded as something of a poor cousin to "theory", but it has in recent years regained value both for its essential empirical role in work on yet un(der)-described endangered languages, and as a key methodological aspect of new developments in the subfield of linguistic typology. *Documentation* has entered linguistics under guise of a specific subfield of "documentary linguistics" whose development and standardization has largely been driven by the accelerated development of new

---

2.   The constructs to be presented here have been amply discussed in recent years within a network of researchers including the University of Lyon research team LED TDR (Langues En Danger: Terrain Documentation and description, Revitalization); its associates in the 3L Consortium (Lyon, London-ELDP SOAS, Leiden); and participants in the 3L series of international summer schools.

technologies. New technologies have also allowed the development of standards for *archiving* of this documentation. We can note here the essential role played by major funding programmes (notably VW DOBES and ELDP-SOAS) in the last 20 years in setting documentation and archiving norms (Gippert, Himmelmann & Mosel 2006).[3] As for the *revitalization* link which has been added to the chain (see Grenoble & Whaley 2006), its status remains better recognized in sociolinguistics and anthropological linguistics than in purely linguistic academic circles, where it may sometimes not even be acknowledged, or, when it is, may be consigned to a lesser-valued realm of "applied linguistics" or "social service", as we examine in more detail below.[4]

The second element of clarification we will propose considers the relation of linguists to the speakers of the endangered languages with whom they work, and to the linguistic communities to which those speakers belong. This particular topic still finds limited space in discussions of ethics in the field of endangered languages, where discussions are more oriented to legal aspects of intellectual property rights and formal definitions of informed consent, than to the nature of the human interactions in the process of collecting data; reflecting the legalistic approach typical of literate societies. Here again, a formula will be proposed to capture succinctly the evolution, over the second half of the 20th century, of a line of thinking preoccupied with the issue of power relations between *researchers* (from the academic sphere in general) and *researched* in the field.

The proposed formula is the simple schema of 'fieldwork *ON, FOR, WITH* and *BY*' as 'fieldwork ON (a language), FOR (a community), WITH and BY (speakers)'. This formula is actually an adaptation and extension of a proposition originally made by Cameron (1992) from the field of sociolinguistics, imported into the field of endangered languages in Craig (1993). It spells a progression from fieldwork ON a language (the ideological schema of the 1950s, of field research using native 'informants') to fieldwork FOR the linguistic community (the ideological schema of the time of civil rights movements, of the *engaged linguist* in

---

3. Several programmes were established around the turn of the century to foster work on endangered languages. The *Dokumentation bedrohter Sprachen* (DoBeS) programme, funded by the Volkswagen foundation of Germany, was based at the Max-Planck Institute for Psycholinguistics, Nijmegen; the Hans Rausing Endangered Language Project (HRELP) was established at SOAS, University of London, in 2002; and the Documentation of Endangered Languages (DEL) program was funded by a consortium of US foundations (the National Science Foundation, the National Endowment for the Humanities and the Smithsonian Institute). It was on this funding base that a new subfield of linguistics named "documentary linguistics" (Himmelman 1998, 2006) took shape.

4. We stress that we do not endorse the positioning of applied research as lower in a hierarchy of status and value; rather, we are critically drawing attention to a common (mis-)conception.

defense of linguistic rights of minorities, taking an advocacy role in "speaking for" the speakers and the communities); to fieldwork WITH the speakers (the ideological schema that emerged in the 1990s, of empowerment, collaborative research and action research (as exemplified in Craig 1992); to which was added, more recently, the final step of the ideal of fieldwork BY speakers trained to be the linguists of their own languages for and within their communities.

This final addition to the schema was expressly requested at an international conference on Amazonian linguistics in Brazil attended by Amazonian indigenous leaders (Grinevald 2000), echoing demands of speakers of Mayan languages (Grinevald 2002, 2006; England 2003), at the commencement of training programs for native speakers like the one at the University of Texas described in Woodbury & England (2004). It brings with it a radical rethinking of the role of linguists in the field, with both synergies and contradictions in terms of expectations in and by the academic sphere, and expectations of the concerned communities of endangered language speakers. As suggested by Costa (p.c.) this is rich material for the further exploration of the links between linguistic rights, diversity, grassroots self-organization, language policy and educational policy and practice. This nexus of issues is outside our current scope but is addressed by other contributions to this volume.

Finally, we would like to offer as last element of clarification a visualization of what Grinevald and Bert (2014) have called the *Spheres of Ideologies* (see Figure 1) within which the different aspects of ELDR practices operate, are conceptualized and are evaluated. For the purpose at hand here, the academic sphere is represented as standing "autonomously", next to and outside the "real world", as represented in the common expression "the ivory tower", or presupposed in the injunction often addressed to the present authors to "not mix science/academia and politics".
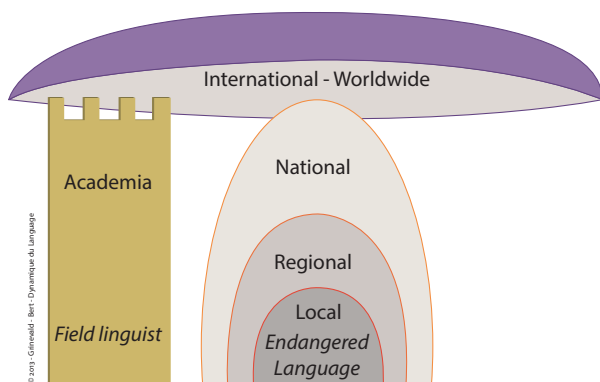

Figure 1. Spheres of ideologies

At the core of the prevailing ideology of the academic sphere is its own sense of its research (and teaching) mission, traditionally articulated as a notion of "pure science" demanding a distancing from what is perceived as the messiness of the "real world." We argue below that this ideological stance is not neutral, as it holds itself to be, but in fact validates relations of unequal exchange, as a consequence of which it is increasingly challenged in many parts of the developing world.

The world "out there" is represented in Figure 1 as a nested arrangement of four stacked spheres of international, national, regional and local scope, from a world-encompassing domain to the very local sphere within which fieldwork takes place on the ground. The local sphere where fieldwork takes place is the sphere where linguists encounter the languages they describe, document, and, in some circumstances, help revitalize. It is the sphere of contact with a linguistic community, and of face-to-face interactions with speakers and all other actors involved in some way with the fate of the endangered languages.

A major issue for ELD researchers is how to deal with the great variety of speakers that one is most likely to encounter in endangered language situations, who not only vary in their levels of pragmatic and grammatical competence as well as lexical knowledge, but also in their attitudes toward the language, as well as their interest in sharing their knowledge. These are some of the basic ingredients that are being considered for the elaboration of a typology of speakers of endangered languages, as proposed in Grinevald and Bert (2010a, b) following up on pioneering work by Dorian (1982). This variety of speakers places on researchers complex demands for interpersonal skills in developing field methods appropriate for local circumstances, a key ingredient of knowledge as practice/ practice as knowledge. Finally, one major issue field linguists must reckon with in the field is the level of consciousness and politicization of the community at large surrounding language issues, including loss of vitality of the language and concern about it.

The national and regional spheres are delimited by constitutions, laws and decrees, and their linked ideologies as expressed in language planning and educational policies, with possible contradictions or even conflicts between national and regional levels. The international and global spheres are the spheres both of forces of globalization that contribute to the loss of language diversity, and of well-established organizations – NGOs or other bodies with diverse philosophies that support or run specific development programs – engaging in the promotion of local endangered languages, as part of larger concerns, such as human rights, indigenous rights, or protection of the environment. This is the case for various branches of the United Nations, such as UNESCO (Minasyan 2014) and its efforts in the past decade on behalf of the protection of intangible world heritages of the world, including languages. What is striking, seen from the field, at least in many

places in Latin America, is the rapidly increasing awareness on the ground, in the local sphere, of the declarations of such international entities.

Our attempt to visually represent spheres and loci of different ideologies regarding endangered languages makes for a rather flat two dimensional schema. The reality in the field, however, is that endangered languages and their communities can either become focal points of attention or lose that attention, accounting for great variations in the ease or difficulty encountered while carrying out a project, eventuating in changes of circumstances very difficult to read and interpret on the spot, even if they can be elucidated later, with hindsight. This instability of "the field" constitutes one of the major elements of risk in such projects for academics who must plan their field trips well in advance and generally from a distance, in order to satisfy the demands of funding agencies and to meet other work commitments. The requirements of funding agencies and universities for rigorous and exhaustive planning, and efficiency demonstrated by quantified results and scientific outputs, often clash dramatically with the realities of the field, creating a gap of divergent interests and ideological views. This gap, which is a source of constant tension for any project of (sustainable) development, is another issue to be cast below in the North-South perspective already mentioned. Doing fieldwork, in our view, in essence consists in a constant back-and-forth between the academic sphere and the local sphere with its diverse links to the other spheres. It is this dialogic and dynamic relational activity, mixing collaboration, contradiction and, sometimes, conflict that we call *knowledge as practice/practice as knowledge*.

We have emphasized in our presentation of the issues so far the multiple scientific and non-scientific contexts that inform the practice and the theory of ELDR. We have also stressed changing conceptions of the roles of ELDR researcher *and* researched. In the following section, we extend the conceptual map of ELDR to encompass its situatedness in global inequalities of wealth and power, both between academe and community, and between different sub-groups of the academe.

3.   North-South relations: Unequal exchange between academics and communities

In this section we focus on relationships between researchers and language communities, framing this in terms of an assumption that the former are from the "North", and the latter from the "South". An important effect of the ideology in which languages are viewed as databanks floating free of the communities that speak them is to validate the notion that language documentation and language

revitalization can and should be regarded as entirely separate enterprises, the former being a dispassionate and value-free scientific activity, and the latter a form of non-scientific "linguistic social work" (Newman 2003: 8). Indeed, major funding programmes for ELDR have explicitly excluded language revitalization from documentation research support. In this view, the relationship between the vitality *vs.* mortality of (respectively) languages, communities and individuals is a monocausal one-way street, in which (once again) communities and individual speakers are merely vehicles for language survival or language death. A language is usually considered dead when the last individual speaker from the original language community passes away. At this point, from the 3rd person perspective of language-as-object, the individual and the community cease to be matters of interest.

In the complex real world, in which there is a dynamic interplay of social, cultural and linguistic processes and practices, this simplistic assumption is contradicted by evidence that language survival can be critical to the life chances of individual members of minority communities, and hence to the viability of the community over time. As Michael Chandler and his colleagues have demonstrated, there is a non-figurative life and death quality to language preservation and revitalization, in relation to individual and community health and wellbeing. Chandler and his colleagues investigated adolescent suicide rates in Canadian First Nations communities, finding that the variable with the single strongest predictive value of *low* rates of suicide, when comparing different bands, was language vitality (Chandler et al. 2003; Hallett, Chandler & Lalonde 2007). These authors' use of the trope of individual, cultural and community *death by language* is a useful corrective to the reifying metaphor of *language death,* the latter being cast as an event that can be conceptually corralled as being of concern principally to the scientific community.

The closely inter-related language-as-object and language-as-databank conceptualizations are constructed from the specific perspective of the scientific community. These conceptualizations not only reify language and divorce language *practices* from the community, they also alienate the language from the community as a form of *symbolic and cultural capital* (Bourdieu 2010). For the communities and their members, the loss of this symbolic cultural capital usually has mainly negative material consequences, such as those documented by Chandler and his colleagues. For the scientist, however, language-as-databank is a form of capital that, once "put to work", can yield symbolic and material benefits in terms of continuing success in grant applications, citations and promotions. Analyzed in this fashion, while the combination of language documentation with language maintenance and revitalization can result in a "win-win" situation, in which there is mutual benefit to both researcher and community, language documentation in

the absence of language maintenance and revitalization is an *unequal exchange* (Emmanuel 1972) in which the benefit to the researcher is not accompanied by an equivalent benefit to the community, and may in fact result in disbenefits.

4.   North-South relations: Unequal exchange inside the academe

Unequal exchange is not confined to the symbolic economic relations between endangered language communities and researchers; it can also be seen in the relations between different groups in the scientific community. Like its geo-political twin the "international community", the term "scientific community" serves to mask dramatic inequalities of power and resource between different members of this *imagined community* (Anderson 2006), and to confer an aura of beneficent disinterest on relations between them. "Northern" linguists receive disproportionately greater funding than "Southern" linguists; they generally enjoy better research infrastructures; have networks that help them update their theoretical, methodological and technical knowledge and skills; and are more likely to have publications in international (Anglophone) journals. These facts generate an ideological construction of *unequal competence*, to the further disbenefit of researchers from the "South". The beneficiaries, linguists from the "North", too frequently seek to maximize their material advantage by laying claim to the intellectual territory of ELDR and its technical wherewithal; by acting as gatekeepers to funding; and by asserting the principle that "data are for everyone", which sometimes boils down to "let me have your data".

   We recognize that many linguists feel deeply affronted by suggestions that their discipline is complicit in the reproduction of unequal relations of power and resources. We suggest that that this is because the objectivist view of science, one manifestation of which is the conceptualization of language-as-object, encourages the elevation of "value neutrality" into a value itself. In our view, on the contrary, the ideology of detached and objective science is best viewed as a cultural model widely adhered to by linguists (amongst others) as a professional body. It underpins the notion of a disinterested scientific community, whose only goal is the pursuit of knowledge, and elides real differences in power, influence and resources between different sub-groups of the imagined scientific community. From such a perspective, even to discuss the politics of linguistic research amounts to a violation of professional norms. We consider, on the contrary, that airing of these issues will both increase awareness of issues of ethics and social responsibility, and contribute to transparency in research practice. Lest it be thought that our critical remarks are fantasies spun out of the web of our own ideological perspective, we

buttress them by providing three brief, anonymized case studies, based on events attested by the authors and personally known colleagues.

Case study 1: Does "data collection" count as research?

A faculty member in linguistics at a well-known and reputable North European university, with no personal ELDR experience, sought data on an Amazonian language. He agreed to fund from his research grant a field trip by two Latin American researchers, one based at the same university, and one based in Latin America, on a topic of mutual interest. The Latin American researchers were an anthropologist and a full professor in linguistics, who between them had more than 25 years field work experience with the community. The Latin American researchers were also familiar with, and experienced in using, the kinds of elicitation tools that the Northern researcher asked them to employ.

After the return of the Europe-based Latin American researcher, during a discussion of publication, the European researcher claimed that the Latin American researchers were working for him, as "assistant" and "facilitator", even though neither was paid a salary by him. He asserted that they had no authorship rights because they had merely "collected data" to which he (because he was funding the field travel costs) had exclusive rights. When challenged, he also claimed that the Latin American researchers lacked adequate linguistic knowledge to plan the research (although they had published internationally on the topic). The Latin American researchers refused to hand over the data, and rejected payment of their field travel costs.

We submit that this case study exemplifies a number of the points we have made above. In particular, the notion that "data collection" is not research, but a merely mechanical activity that requires no particular knowledge or experience, betrays a failure to comprehend the complex activity that we have dubbed "knowledge as practice/practice as knowledge". This failure can be (and in this case, was, institutional) as well as individual: the university, although made aware of the dispute, declined to take any position. We note, further, that the attempted appropriation of the right unilaterally to determine rights of authorship is contrary to internationally accepted ethical norms, as laid down in for example International Committee of Medical Journal Editors (2014).

Case study 2: Are field sites "open territory"?

A faculty member in linguistics at a well-known and reputable North American university, with significant field ELDR experience, planned and initiated a large-

scale typological project, for which graduate students were to be dispatched to various communities in a Latin American country to collect data. One of the graduate students was sent to a community with no prior contact with a linguist who is a national of the Latin American country (known to the "Northern" faculty member), who had been working in and with that community for many years. The Latin American linguist had already collected data on the particular topic under research, but this fact was not made known to the graduate student. When the Latin American linguist next made a field work visit to the community, community members expressed their unhappiness and displeasure at the North American graduate student's visit, demanding to know what the Latin American linguist's role was in that field trip. The reason for the community members' negative reaction was that the graduate student had received, from their home institution, insufficient background information and had not been adequately briefed in issues of sensitivity and ethics.

We submit that this case study exemplifies a framing of the relationship between researchers and community in which "the field" is viewed exclusively as a "data source", rather than a site of engagement between researchers from different teams and members of the community. It is common, to the point of being almost unavoidable, that long-term engagement with a community not only enables the development of a relationship of trust, but also brings with it vulnerabilities in which perceived violations of trust can have enduring negative consequences for the researcher who is in it for the long haul. The immediate cause for the problems in this case was a failure of communication, but the background to this failure was a lack of recognition by the "Northern" researcher both of the vulnerabilities of the "Southern" researcher and of the importance of establishing exactly what the Southern researcher already knew about the topic under investigation; the latter perhaps symptomatic of a widespread attitude that linguistic data are independent of social relationships. Our argument is that communities as "field sites", while not being the "property" of any individual or team, cannot be separated from communities as societies in relationship to researchers, in which some relationships are more enduring and vulnerable than others.

Case study 3: To what extent, and when, should data be in the public domain?

A faculty member in linguistics at a well-known and reputable North European university, with no personal ELDR experience, initiated a large-scale, collaborative typological project on a theme overlapping with that of a previous project based in a different country of the North. The funding for the new project, which

was conceived together with a linguist in the second country of the North, was acquired on the assumption that data collected for the previous project on indigenous languages of South America, mostly by PhD students and postdoctoral researchers, would be made freely available for the new project, with a view to the leaders of the new project publishing the analyses. The North European linguist, who had not discussed this assumption with the data collection team of the first project, asked at an initial project meeting (for the new project) for the already existing data from the previous project to be put at his disposal for him to analyze and publish. He was surprised and annoyed when the researchers who had carried out and supervised the prior field work (who, unlike the North European linguist, specialized in the regional languages under investigation) were not willing to agree to this.

This case study does not, perhaps, so directly implicate North-South relations as the previous two, although some of the field researchers were "Southern". It does, however, exemplify the major and contentious issue of open access to, and open archiving of, data. The "piggy-backing" of new projects on previous projects, with cumulative re-analyses of data, is a frequent and often productive research strategy. Open archiving is also increasingly mandated by research funding agencies. It would seem at first sight that this can bring only benefits to the scientific community, and we do not wish to downplay these benefits. However, there are also potential disbenefits to researchers, particularly those from the "South" who have less experience of international (Anglophone) publication than "Northern" researchers and who are less likely to be the leaders of international consortia. We would also draw attention to the potential disbenefits for communities of "absolutist" interpretations of open data archiving and open data access. There are genuine reasons why communities, as well as primary field researchers, may wish to restrict access to certain genres of linguistic data, including those that are restricted within the community itself. This issue becomes even more salient in the case where the primary field researcher is also a native speaker. Researchers from the "North" should be aware of these issues from the start when planning new projects, and be sensitive and flexible enough to negotiate mutually beneficial agreements with communities and with "Southern" colleagues.

General discussion

The three case studies that we have briefly presented here are attested by our own experience, together with that of trusted colleagues. They are representative of a number of other cases that have come to our attention, and that are also reliably attested. They should therefore not be dismissed as exceptional outliers. Rather,

such cases are best understood as manifestations of a systemic problem that is tightly bound up with particular notions of what is, and is not, science; of who is to be considered a qualified scientist; and of who has intellectual property rights over data. This systemic problem is deeply rooted in the objectivist view of science as an activity that is not only value-free, but essentially context-free.

We referred to and analyzed the devalorization of local knowledge in Section 1. Here we emphasize that this devalorization is all too often mapped onto North-South relations, in cases where the work of researchers from the "South" may be dismissed as old-fashioned or lacking in technical sophistication, without a complementary and reflexive understanding of the lack of local knowledge on the part of the researcher from the "North". Capacity building is all too often viewed as a one-way North-to-South transfer of knowledge, with data and intellectual capital accruing unequally to the "North." The presupposition of unequal distribution of knowledge and expertise, conceptualized in terms of a hierarchy of skills and knowledge, forms a seamless ideological join with real world inequalities of power and resources, in which those with access to large scale funding and experience of international publication take it as their right to determine the conditions of "collaboration." It is appropriate to characterize such attitudes and assumptions as neo-colonialist (we do not claim these to be universal, but simply assert that they exist, as we have documented in the first two case studies).

Concluding reflections on theory, methodology, goals and values

ELDR is not, and should not be seen as, a world unto itself. It takes place within a complex array of contexts, including relations of global socio-economic inequality, North-South relations within the academe, and the existential predicaments of indigenous minority communities. These communities face multiple threats to their traditional ways of life, including loss of environment, depletion of resources and the hostility of neighboring communities, as well as the increasing penetration of the dominant language and culture. Rather than viewing them solely as "vanishing cultures" to be archived and memorialized, it is more useful and more ethical to view them as communities challenged by complex cultural dynamics, and to assist them in the formation of strategies that will secure their future as equal co-participants in national development. ELDR should be situated in the realities of dynamic cultural change, the adaptation of tradition to new circumstances and the renewal of culture as a mode of participation in a changing world. It is, on this view, an integral part of an overall strategy encompassing community wellbeing in health, education, economy and environment. Implementing such a strategy requires the participation of scientists and professional practitioners

from many different disciplines, and the search for effective ways of complementing scientifically useful knowledge with practically useful, community guided interventions.

Unlike many social sciences (including anthropology and cultural psychology), linguistics does not traditionally value reflexivity as an inherent part of the research process.[5] Indeed, many linguists would deny that linguistics is a social science, preferring to emulate the supposedly objective (and objectifying), disengaged stance of the natural sciences. In relation to ELDR, this leads to a view in which socio-political questions relating to revitalization, community engagement and empowerment are compartmentalized as questions of "mere application", or of individual opinion and responsibility. We maintain, on the contrary, that *linguistics is a social science*, and its practice in ELDR is not of one neutral inquiry by a disinterested observer, but a social practice with social consequences.

We have criticized the view that languages are primarily "science heritage" resources, and only secondarily vital to the survival of indigenous communities. We have also argued that the cultural model of objectivist and "value-free" science encourages the self-distancing of linguistic ELD researchers both from issues of the distribution of knowledge and power, and from engagement in practices of language revitalization. It also, at least in part, underpins unequal exchange relations between linguistics researchers from the global "North" and the global "South". We would advocate a different model in which research is a collaborative learning enterprise, involving mutual knowledge transfer to mutual benefit, analogously with the WITH and BY paradigms we discussed above in relation to researchers' relations with endangered language communities. Collaborative relations can and should involve two-way capacity building and scientific empowerment (including empowerment OF, WITH and BY native speakers: Cabral, Sampaio & Silva Sinha, this volume; Woodbury & England 2004).

Although we do consider the dissemination throughout the research community (including indigenous researchers) of knowledge and skill in using up-to-date technology for recording, annotation, description and archiving of data to be of great importance, we would also argue that this should be complemented by the dissemination of an understanding of the relevant methodological and

---

5.   Disciplinary self-reflection in anthropology (e.g. Clifford & Marcus 1986) and cultural psychology (e.g. Gergen et al. 1996) has long included discussions of power inequalities in both researcher-researched and North-South collaboration relations, in a global context; and has been contextualized too by reflexive debate on the cross-cultural appropriateness of methods and theories (Cole et al. 1971). Such reflection was for a long time virtually absent in linguistics, although it is now becoming part of ELDR debates (Austin & Sallabank 2014). We can hope that this attention to reflexivity will take hold in the wider discipline of linguistics.

ethical principles of qualitative research and of research in indigenous cultures. We would therefore advocate that just as much attention is paid, in ELDR education and training, to the socio-cultural context and dimensions of ELDR as to its linguistic-theoretical and technical aspects. ELDR is, or should be, an interdisciplinary activity, and its theory, education and practice should reflect this. The theory, methodology and practice of ELDR needs to be informed as much by the reflexive and qualitative stance of disciplines such as anthropology, as by the experimentalist and quantitative stance of disciplines such as cognitive psychology or corpus linguistics. ELDR education and training should reflect the dynamics and complexities of knowledge as practice/practice as knowledge, only part of which is knowledge of how to use technical equipment and software tools.

We would advocate, too, that university ethics courses and ethics committees should adopt and enforce principles and criteria relating to community rights, community engagement, community intellectual property and principles of collaborative research, in addition to the usual focus on individual informed consent. Novice ELDR researchers should be presented with examples of best practice in combining documentation and archiving with language preservation and revitalization. Above all, our message is that the critical discussion of the issues we have raised should not be seen as a distraction from "business as usual", but as an integral part of the endangered languages research landscape.

References

Anderson, B. 2006. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. London: Verso Books.

Austin, P. & Sallabank, J. (eds). 2010. *The Cambridge Handbook of Endangered Languages*, 45–65. Cambridge: CUP.

Brenzinger, B. (ed.). 2006. *Endangered Languages*. Berlin: Mouton de Gruyter.

Bourdieu, P. 2010. The forms of capital. In *Cultural Theory: An Anthology*, I. Szeman & T. Kaposy (eds), 81–93. New York NY: John Wiley & Sons.

Cameron, D., Frazer, E., Harvey, P., Rampton, B. & Richardson, K. 1992. Introduction. In *Researching Language: Issues of Power and Method*, 1–28. London: Routledge.

Chandler, M.J., Lalonde, C., Sokol, B., Hallett, D. & Marcia, J. 2003. Personal persistence, identity development, and suicide: A study of native and non-native North American adolescents. *Monographs of the Society for Research in Child Development* 68(2): i–138.

Clifford, J. & Marcus, G.E. 1986. *Writing Culture: The Poetics and Politics of Ethnography*. Berkeley CA: University of California Press.

Cole, M., Gay, J., Glick, J. & Sharp, D. 1971. *The Cultural Context of Learning and Thinking: An Exploration in Experimental Anthropology*. New York NY: Basic Books.

Craig, C. 1992. Miss Nora, rescuer of the Rama language: A story of power and empowerment. In *Locating Power, Proceedings of the Second Berkeley Women and Language Conference*, K. Hall, M. Bucholtz & B. Moonwomon (eds), 80–89. Berkeley.

Craig, C. 1993. Commentary on: Ethics, advocacy and empowerment: Issues in methods in researching language. *Journal of Language and Communication* 13(2): 81–94. doi:10.1016/0271-5309(93)90001-4

Dawkins, R. 1976. *The Selfish Gene*. Oxford: OUP.

Dorian, N.C. 1982. Defining the speech community to include its working margins. In *Sociolinguistic Variation in Speech Communities*, S. Romaine (ed.), 25–33. London: Edward Arnold.

Emmanuel, A., Bettelheim, C. & Pearce, B. 1972. *Unequal Exchange: A Study of the Imperialism of Trade*. New York NY: Monthly Review Press.

Everett, D. 2005. Cultural constraints on grammar and cognition in Pirahã. *Current Anthropology* 46: 621–646. doi:10.1086/431525

Everett, D. 2009. Pirahã culture and grammar: A response to some criticisms. *Language* 85: 405–422. doi:10.1353/lan.0.0104

Gergen, K.J., Gulerc, A., Lock, A. & Misra, G. 1996. Psychological science in cultural context. *American Psychologist* 51: 496–503. doi:10.1037/0003-066X.51.5.496

Gippert, J., Himmelman, N. & Mosel, U. (eds). 2006. *Essentials of Language Documentation* [Trends in Linguistics: Studies and Monographs 178]. Berlin: Mouton de Gruyter. doi:10.1515/9783110197730

Grenoble, L. & Whaley, L. 2006. *Saving Languages: An Introduction to Language Revitalization*. Cambridge: CUP.

Grinevald, C. 2000. Los lingüistas frente a las lenguas indígenas. In *As Línguas Amazônicas Hoje*, F. Queixalos & O. Renault-Lescure (eds), 35–53. Saõ Paolo: IRD MPEG Instituto Socioambiental,

Grinevald, C. 2002. Linguistique et langues mayas du Guatemala. *Faits de Langues. Meso-Amerique, Caraibes, Amazonie* 20(1): 17–27.

Grinevald, C. 2006. A view from the field: An Amerindian view, worrying about ethics and wondering about informed consent. In *Lesser Known Languages in South Asia: Status and Policies, Case Studies and Applications of Information Technology*, A. Saxena & L. Borin (eds). Berlin: Mouton de Gruyter.

Grinevald, C. & Bert, M. (eds). 2010a. *Linguistique de Terrain sur Langues en Danger: Locuteurs et Linguiste. Faits de Langues*, 35–36. Paris: Ophrys.

Grinevald, C. & Bert, M. 2010b. Speakers and community. In Austin & Sallabank (eds), 45–65.

Grinevald, C. & Bert, M. 2014. Whose ideology, when and where: Revitalization of Rama (Nicaragua) and Francoprovençal (France). In Austin & Sallabank (eds), 357–385.

Hallett, D., Chandler, M. & Lalonde, C. 2007. Aboriginal language knowledge and youth suicide. *Cognitive Development* 22: 392–399. doi:10.1016/j.cogdev.2007.02.001

Himmelmann, N.P. 1998. Documentary and descriptive linguistics. *Linguistics* 36: 161–196. doi:10.1515/ling.1998.36.1.161

Himmelmann, N.P. 2006. Language documentation: What is it and what is it good for? In Gippert, Himmelman & Mosel (eds), 1–30. 10.1515/9783110197730

International Committee of Medical Journal Editors. 2014. Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals, 29 July 2015. <http://www.icmje.org/icmje-recommendations.pdf>

Minasyan, A. 2014. UN discourse on linguistic diversity and multilingualism in the 2000s: Actor analysis, ideological foundations, and instrumental functions. In Austin & Sallabank (eds), 385–406.

Nevins, A., Pesetsky, D. & Rodrigues, C. 2009. Pirahã exceptionality: A reassessment. *Language* 85: 355–404. doi: 10.1353/lan.0.0107

Newman, P. 2003. The endangered languages issue as a hopeless cause. In *Language Death and Language Maintenance* [Current Issues in Linguistic Theory 240], M. Janse & S. Tol (eds), 1–13. Amsterdam: John Benjamins. doi: 10.1075/cilt.240.03new

Sharifian, F. (ed.). 2015. *The Routledge Handbook of Language and Culture*. Abingdon: Routledge.

Woodbury, A. & England, N. 2004. Training speakers of indigenous languages of Latin America at a US university. In *Language Documentation and Description*, 2, P.K. Austin (ed.), 122–139. London: SOAS.

Indigenous language policies in Brazil

Training indigenous people as teachers and researchers

Ana Suelly Arruda Câmara Cabral, Wany Bernardete de Araujo
Sampaio and Vera Da Silva Sinha
University of Brasília / Federal University of Rondônia /
University of East Anglia

In this chapter we outline the historical background of Brazilian language policies that are meant to be supportive of Brazilian indigenous languages and discuss some positive and negative impacts of national programmes developed under these policies. We single out the official programmes relating to indigenous education, language planning and language description and documentation, and ways in which they might help to assure the survival and vitality of Brazil's indigenous languages. We also draw attention to official support for directing these positive initiatives towards an effective strengthening of indigenous languages and cultures in collaboration with indigenous peoples and in the context of self-determination and empowerment.

Keywords: documentation, indigenous education, indigenous languages of Brazil, language description, language planning, linguistic policies

1.   Introduction

Brazil is the South American country with the richest linguistic diversity. Excluding Libras, other sign languages and Brazilian Portuguese, there are approximately 220 languages, comprising immigrant languages and indigenous languages and their variants (Rodrigues 2012). We have to add the possible existence of at least 40 more indigenous languages which are spoken by isolated communities who have not yet experienced any contact with present day outsiders that have been ignored in statistics concerning extant Brazilian indigenous languages. Most of the indigenous languages in Brazil are seriously endangered, either because their use presents symptoms of progressive decline, such as reduction in the contexts

of use, or because they are no longer being transmitted to new generations. This is a consequence of many years of assimilationist policies promoted by the Brazilian Federal government, whose goal until recently had been the integration of the so-called minority indigenous communities into the dominant Brazilian society, discouraging and suppressing thousands of years of linguistic diversity and indigenous knowledge.

The last thirty years, however, have seen remarkable advances in the scenario of linguistic policies concerning indigenous languages. These include an increase in linguistic description and documentation of these languages; an increasing number of doctoral and master's dissertations on indigenous languages; increased financial support from governmental agencies for linguistic studies of indigenous languages. Most importantly, policies have been implemented for the promotion of indigenous intercultural education, as well as the inclusion of indigenous students in undergraduate and graduate programmes by means of quota policies. A new program, The National Inventory of Indigenous Languages (an initiative of the Ministry of Culture) also reflects a positive governmental approach to safeguarding indigenous languages and cultures.

In this chapter, we shall (i) briefly outline the historical background of Brazilian language policies meant to be supportive of indigenous languages; (ii) discuss some positive and negative impacts of national programmes developed under these policies. Our main focus here is on the official programmes concerning indigenous education, language planning and language description and documentation, and how they may help to assure the survival and vitality of indigenous languages of Brazil. We shall call attention to the kind of official political support necessary to direct these positive initiatives towards a more effective strengthening of indigenous languages and cultures in partnership with the indigenous peoples in their movements for self-determination and empowerment.

2. A brief account of the last thirty years of indigenous languages research and documentation

The linguistic description and documentation of Brazilian indigenous languages started some thirty years ago. In 1986, the Brazilian linguist Aryon Dall'Igna Rodrigues, in association with Dr. Marisa Cassim, a technical adviser of the National Council for Scientific and Technological Development (CNPq), created the *Programa de Pesquisa Científica das Línguas Indígenas Brasileiras* (PPCLIB: Scientific Research Program of Brazilian Indigenous Languages), the first official program supporting graduate studies, field work research and documentation of Brazilian indigenous languages. This program was the first step in the legitimization of

public policies recognizing the importance of the linguistic study and documentation of indigenous languages of Brazil. The PPCLIB provided: (i) delivery of four short courses preparing students of linguistics for the scientific study of Brazilian indigenous languages; (ii) the granting of fellowships for graduate studies abroad focusing on the description and documentation of Brazilian indigenous languages; (iii) the granting of fellowships for masters' students in Brazilian universities for the purpose of scientific research and documentation of Brazilian indigenous languages;[1] and (iv) funding for field work on indigenous languages. The proposal setting out the CNPq policies took into consideration the inventory made by Rodrigues (1985, 1986) summarizing the number of extant Brazilian languages, their importance and the urgent need of developing their linguistic study and documentation.One of the most significant research projects benefiting from the PPCLIB, coordinated by the linguist Lucy Seki, aimed at the description and documentation of Xingu languages such as *Trumai, Awetí, Suyá, Kamaiurá* and *Jurúna*. The efforts of Aryon Dall'Igna Rodrigues and Marisa Cassim also secured the appointment of a linguist at the Goeldi Museum in Belem. The PPCLIB became the most important program ever of scientific research on Brazilian indigenous languages.

Although the PPCLIB was suspended in 1990, these initiatives generated fruitful results. Indigenous languages entered the list of scientific fields funded by CNPq; and at the research training agency *CAPES*, indigenous languages became an official field of graduate studies. As a consequence, linguistic studies of Brazilian indigenous languages proliferated in graduate programmes in different parts of Brazil: (the FederalUniversities of Pará, Pernambuco, Alagoas and Santa Catarina, and the University of Brasília.

Another program of fundamental importance in the development of Brazilian policies regarding indigenous languages and cultures was the *Projeto Interação* (Interaction Project) (1983–1989). It was carried out by the *Pró-Memória* National Foundation, a now defunct Department of the *Instituto do Patrimônio*

---

1.   Amongst the positive results of the PPCLIB is the number of students that finished their PhD abroad, between 1989 and 2004: Filomena Sândalo, *A Grammar of Kadiwéu* 1995, University of Pittsburgh, PITT, USA.; Ana Suelly Arruda Câmara Cabral, *Contact Induced Language Change in the Western Amazon: The Non-Genetic Origin of the Kokama Language*, 1995 University of Pittsburgh, PITT, USA. Nilson Gabas Junior, *A Grammar of Karo* (Tupi, Brazil) 1999 University of California System, UC System, USA. Marcia Dâmaso Vieira, *O problema da não-configuracionalidade na língua Asurini do Trocará: um fenômeno derivado da projeção dos argumentos verbais* 1993, University of Arizona, USA. 1993. Marilia Facó Soares, *O suprasegmental em Tikuna e a Teoria Fonológica. Volume I: Aspectos da Sintaxe em Tikuna*, Universidade Estadual de Campinas, UNICAMP, Brasil. 1984–1992; Luciana Raccanelo Storto, *Aspects of a Karitiana Grammar* 1999. MIT, USA. (See also Seki 1999).

Histórico e Artístico Nacional (Institute of National Historical and Artistic Heritage), founded during the transition from twenty years of military dictatorship to democracy, inaugurating the first official program promoting: (i) the interface of culture and education; (ii) educational projects oriented to the different Brazilian sociocultural contexts; (iii) focusing on cultural and linguistic diversity; (iv) inspiring the dissemination of intercultural knowledge, and (v) funding civil society as well as governmental organizations developing projects in this field (Quintas 1996).

Two main projects benefiting indigenous groups were supported by the *Projeto Interação*: the Authorship Project of the Indigenous People of the Acre state (Cabral et al. 1987), and the Tikúna Project; both of them charged with bilingual education and with the documentation of indigenous languages and cultures. Very importantly, it was the indigenous communities themselves who were the the main agents in the educational and documentation processes. Since in this project the indigenous students started having access to the writing systems of their own languages, they also started producing a written indigenous literature based on their myths and other traditional native knowledge. For the first time in Brazil, indigenous teachers were contracted by State governmental education departments, and indigenous languages were officially recognized as the legitimate languages of instruction in indigenous schools. The *Projeto Interação* financed some of the first books mainly authored by indigenous writers. Amongst these were *Torudü'ügu,* a bilingual reader of Tikuna myths, translated into Portuguese, and a significant number of other books written by indigenous students of the Acre project.

Projeto Interação, with its innovative mission and political orientation, became an important reference during the 1980s; it had a positive impact in the writing of the 1988 Brazilian Constitution, and subsequent governmental policies regarding indigenous cultural, linguistic and education rights. The 1988 Constitution guaranteed for the first time the right of indigenous peoples to primary education in their own language. It charged the official educational system with providing to indigenous communities the recovery of their historical memory, the reaffirmation of their ethnic identities and the valorization of their indigenous languages and sciences. It also prescribed the recognition of indigenous peoples' forms of social organization, customs, languages and traditions; and recognized their natural rights to the lands traditionally occupied by them, with responsibility for the demarcation of their lands and the protection of their heritages being assigned to the Federal Government.

These constitutional measures were reflected in the 1999 Education legislation (1999) which has inspired and validated new official policies and educational

programmes responsible for the strengthening of indigenous languages and cultures. New addenda to this Law have been issued, further improving indigenous educational policies. Monserrat (2000) observes in 1999 there were 3,127 indigenous students and 1,673 indigenous schools in the indigenous reservations. A report produced by the Ministry of Education in 1998 identified, as highlighted by Monserrat (2000: 142), the absence, with some exceptions, of the use of the indigenous languages as a specific discipline, as the languages of instruction of other disciplines, and as languages of literacy programmes. Among the exceptions, we can cite the indigenous Tikuna, Roraima, Yanomami, Kaiová and Guaranis schools of Acre where the native language was the main language of

During the 1990s, Brazilian policies developed through the provision of programmes aiming to improve indigenous primary schooling and indigenous teachers education. The responsibility for indigenous education, which previously rested with the *Fundação Nacional do Índio* (National Indian Foundation, a federal organ), became the responsibility of each State with an indigenous population. State Departments of Education started developing educational programmes to prepare native teachers in indigenous reservation schools. The pioneering programmes coordinated by civil organizations, such as the Acre experience of indigenous authorship, continued to be an important reference for this process, as well as the ideas promoted by *Projeto Interação*.

Around the turn of the century, further significant changes took place in the field of indigenous education in Brazil. The experience of transferring indigenous education to the governmental sphere required increased specialization and the teaching of native languages required the training of native speakers as teachers.

In 2004, the *Secretaria de Educação Continuada, Alfabetização e Diversidade* (SECAD: Department of Continuing Education, Literacy and Diversity) was established within the Ministry of Education and Culture. This Department brought together, for the first time, themes such as field education, environmental education, indigenous education, ethnic diversity and racial issues, previously distributed amongst different ministerial sectors. From 2011 SECAD has included in its mission social inclusion, with the addition of this to its title. (Department of Continuing Education, Diversity, Diversity and Inclusion: SECADI). The main responsibilities of SECADI with respect to the indigenous populations are:

1. Initial and continuing training of indigenous teachers at intermediate level. These programmes have an average duration of five years, leading to a formal Diploma qualification, the *Magisterium*. They are delivered mainly by residential short courses, in which indigenous teachers leave their villages for a month, participating in joint activities at a training center, and stages of independent studies, research and reflection on teaching practice in the

villages. The Education Ministry provides technical and financial support for the courses.

2. Higher Education of indigenous teachers (Intercultural undergraduate courses). The main objective is to ensure quality in indigenous education and expand the provision of teaching to the secondary school and high school sectors in indigenous communities and territories.

3. Production of didactic materials in indigenous language, bilingual or Portuguese formats. Books, posters, videos, CDs, DVDs and other materials produced by indigenous teachers are published with the financial support of the Education Ministry and distributed to indigenous schools.

4. Political and pedagogical support to school systems for expanding the provision of education in indigenous lands.

5. Promotion of indigenous social autonomy. The MEC develops, in conjunction with FUNAI, training courses for teachers and indigenous leaders; these courses contribute to knowledge of civil and social rights and to the exercise of control over the mechanisms of funding of public education, as well as the implementation of actions and programmes in support of indigenous education.

6. Financial support for the construction, renovation and expansion of indigenous schools.

3.   The implementation of linguistic, cultural and educational policies benefiting indigenous peoples in Brazil

Currently, in Brazil, every Department of Education in the States with indigenous populations provides programmes to indigenous teachers at the high school level; some public universities, in these same states, have Intercultural Indigenous Programmes. There are states that have more than one Intercultural Programme, particularly those with large numbers of indigenous communities, as is the case for the States of Mato Grosso, Amazonas and Pará. The idea is that this degree programme admits graduates of the High School Teacher Training (*High School Magisterium*). Let us take for example the state of Rondônia: in 2004, the High School Magisterium was instituted for all indigenous teachers of Rondônia; at that time, the number of indigenous communities with primary schools was negligible. Initial contact with some indigenous communities took place in late 1979 and early 1980; the first establishment of elementary schools gathered momentum in the late 1990s; most students at High School Magisterium level were between 15 and 20 years old.

Under Brazilian law, indigenous peoples have the right to specific, differentiated, intercultural and bilingual schooling. In summary, the most important policies are set out in:

- The Federal Brazilian Constitution (1988), which guarantees to the indigenous people's rights to their own social organization, customs, languages, beliefs and traditions.
- Presidential Decree no. 26 (1991) – which transfer the coordination of educational activities from the Ministry of Justice (FUNAI) to the Ministry of Education; the implementation of these actions is assigned to States and Municipalities.
- Law no. 9,394 – Law of Guidelines and Bases for National Education (1996), establishing the provision of bilingual and intercultural education as a duty of the State.
- Resolution no. 3 (CEB/CNE, 1999), which establishes national curriculum guidelines and standards for the recognition and operation of indigenous schools.
- The National Guidelines for Indigenous Education (1993), for differentiated, bilingual and multicultural education of indigenous peoples and training of indigenous teachers.
- The National Curriculum Guidelines for Indigenous Schools (1998), which provide insight and guidance on the preparation of indigenous education programmes to meet the aspirations and interests of indigenous communities, with regard to the principles of cultural equity among all sections of Brazilian society, as well as the development and production of teaching materials and indigenous teacher training.
- The Indigenous Teachers Formation Guidelines (2002), whose objectives are to contribute to the creation and implementation of initial and continuing training of indigenous teachers in State education systems, and programmes to meet the demands of indigenous communities.

In respect to the continuing education of indigenous teachers, important policies have been developed by means of governmental programmes such as:

- The "Parameters in Action of the Indigenous Scholar Education Program" (initiated in 2002), which aims to support the "professional development of teachers and experts in education, in coordination with the implementation of the Parameters and National Curriculum Guidelines for Elementary Education, for Indigenous Education and Childhood Education; and also for the implementation of Youth and Adults Education" (p. 05).

– The Action "Indigenous Knowledge in School" (initiated by *Portaria no. 98*, December 2013), whose objectives are:

I   to promote the continuing education of teachers who work in indigenous primary education;

II   to provide instructional and learning resources that respond to the specificities of the community organization, multilingualism and interculturalism underpinning educational projects in indigenous communities;

III to provide aids for curriculum development, development of methodologies and evaluation processes that respond to the specific processes of literacy, numeracy and knowledge of indigenous peoples;

IV   to promote researches that result in the production of bilingual and monolingual didactic and paradidatic materials, according to the sociolinguistic situation and in accordance with the specific needs of indigenous scholar education.

3.1   The consolidation of educational programmes for indigenous communities

A recent project for training indigenous primary school teachers is the *Açaí* Project, which is a modular four year educational project developed by the Rondônia State Department of Education.

The *Açaí* Project involved, between 1998 and 2004, 126 indigenous teachers from 34 ethnic groups, of whom 44 were already teachers in their village school (cf. Gonçalves 2013, 86).[2] One of the hallmarks of this project has been an interdisciplinary conception and the inclusion of professionals with extensive experience in elementary school. As a fundamental practice, beside the inclusion of non-indigenous teachers with extensive experience in key disciplines for training, the programme also included some indigenous teachers who had extensive experience with the wider indigenous issues and indigenous education. The course systematically brought together the same teachers twice a year, so that there was continuity in the application of methods and content.

Given the linguistic diversity represented by 14 indigenous ethnic groups with a preponderance of students of *Mondé* and *Ramaráma* families, the latter were organized as a separate class pursuing the course "Indigenous Languages";

---

2. The project was planned to be developed over five years, totaling 4620 hours, divided into two annual intensive steps of approximately 35 days each, corresponding to four or five weeks with workload around 250 hours per step, totaling, at the end of the course, 2700 hours of class contact time and 1920 hours of homework (at the village school), with the pedagogical support (Gonçalves 2013: 87).

the other students (*Oro, Oronáwa, Aikanã, Kuaza, Kanoê, Tupari Makurap, Jeoromitxi, Nambiquara, Uru-eu-au-au*) were gathered in another "Indigenous Languages" class.

Almost all the students who completed the first round of the *Açaí* Project (1998–2004) subsequently entered the Teacher Training programme "Superior Indigenous Intercultural Magisterium", located in the Ji-Paraná campus of the Federal University of Rondônia. This program was designed by some of the same academics who had designed the *Açaí* Project. In this sense, for indigenous students, the entry into college represented the continuation into a new stage of studies supported by familiarity with the colleagues and teachers.

3.2   The undergraduate programme in Intercultural and Basic Education of the Department of Intercultural Education of the Federal University of Rondônia

A major goal of undergraduate programmes such as that in Intercultural and Basic Education of the Department of Intercultural Education of the Federal University of Rondônia is preparing these students for language teaching in village schools. This is perhaps the biggest challenge for all programmes, for several reasons. The principal one is the lack of established methods of mother tongue teaching in schools; from the initial acquisition of literacy in the mother tongue, the students are expected to "jump" directly to the composition of texts. However, the indigenous teachers do not have sufficient linguistic knowledge of the language and they fully mastered writing in their native language. There is therefore still a lot of work to be done on writing in such a way as to conform to the structure and grammatical principles of each language.

One of the biggest problems of indigenous intercultural programmes is undoubtedly the teaching of indigenous languages. In programmes such as that in Ji-Paraná, at least 8–10 languages are spoken in each class, all belonging to different genetic families, such as *Jeoromitxi, Paiter, Kanoê, Makurap, Oro Win, Tupari, Karitiana, Karipuna, Puruborá and Zoró,* for example. The question then is: how can we appropriately and thoroughly deliver indigenous languages courses in classes with such linguistic diversity? There is great motivation on the part of the students to learn the linguistic analysis of the grammar of their native languages, but how can an intercultural programme respond to the demands of such a diverse reality?

The Ji-Paraná programme has invested in the participation of linguists for teaching phonetics, phonology, morphology and syntax. But the time devoted to these disciplines is inadequate for an effective teaching of the linguistic concepts that would allow students to develop a thorough theoretical knowledge of the

structure of their native languages. There is also the problem that not every linguist who teaches on these courses has experience with language teaching; only a minority are both descriptive linguists and specialists in applied linguistics.

The indigenous students aspire to theoretical-descriptive knowledge of their languages, but the existing grammars of the indigenous languages of Rondônia are almost all written in English. They are useless for the indigenous students, who do not read English. Even the grammars and other descriptive works in Portuguese are extremely technical and difficult to "digest" by indigenous students.

One of the big issues that the Indigenous Intercultural Programmes is focusing on is the extent to which linguists, and their academic research, may contribute to the training of indigenous teachers so that they can develop their language teaching in villages schools. In Rondônia, even in groups in which indigenous languages are still fully transmitted to the new generations there is a competition between Portuguese and the indigenous language, and even though teaching is given in the native language, the Portuguese language comes with a didactic framework which makes it a strong competitor to the indigenous languages. There is a huge lack of written materials in indigenous languages that can be used in the teaching of these languages in the village schools. Some linguists help produce brochures for literacy, books with stories collected from older people, but these materials still have insufficient impact on the daily life of the village schools.

Another important aspect of the problem of indigenous teacher education, at both High School and undergraduate levels, is the degree of vitality of languages, which differs greatly from one language to another. Some languages like *Campé* and *Arikapú* are spoken by a very small number of individuals (in the case of *Campé,* the three remaining speakers live practically without communicating with each other as a result of the geographical distance separating them); other languages like *Puruborá* are no longer spoken, and most of the languages of the *Txapakúra* family are no longer being transmitted to the new generations. This complex reality requires the development of linguistic studies and special teaching methods for each particular case. There is a need for specialists in methods of language teaching not only as a first language, but also as a second language, and a category of language teaching which shares features of a foreign language. And how can we adequately provide programmes in languages that are only partially remembered, but in which the remaining knowledge of the language is of fundamental importance to the communities where these languages were spoken in the past?

In Intercultural Programmes where linguistic diversity is minimal, the problems relating to the teaching of indigenous languages are fewer, as is the case at the Federal University of Minas Gerais*,* where *Maxakali* is the only language fully

spoken; the Federal University of Goiás, in which students are *Xerente, Tapirapé* and *Karajá, Kamayurá* and *Kalapalo* (these two last ones being included only in 2014); the Federal University of Grande Dourados has the privilege of hosting students all speaking languages from Sub-branch I of the Tupi-Guarani family: *Nhandewa, Ymbiá* and *Kaiowá.* However, the programmes located in the Federal Universities of Acre, Amazonas*,* Mato Grosso and Rondônia face a problematic and complex situation of linguistic diversity that seems to have no immediate solution.

It is to be hoped that these experiences, together, can nourish reflections that lead to a new approach to the training of indigenous teachers that actually prepares them to deal with the teaching of native languages in the village schools. Experience has shown that new policies and planning are needed to resolve this impasse.

In a country with such linguistic diversity as we find in Brazil, it is necessary to create Applied Linguistic programmes for the teaching of indigenous languages. In every public University, there are Departments of Linguistics and Applied Linguistics. In some of these Universities there are Departments with research lines emphasizing indigenous languages, like the University of Brasília, the State University of Campinas, the Federal University of Rio de Janeiro, the State University of São Paulo, the Federal University of Goiás*,* the Federal University of Pará, the Federal University of Alagoas, among others; but in none of them do the Applied Linguistics programmes address indigenous languages, even though they teach languages such as English, French, German, Spanish, Italian, Persian, Japanese, Greek and Latin. None of the more than 180 Brazilian indigenous languages are taught.

3.3   The challenge of including indigenous students in the Linguistics Graduate Programmes in Brazil

The University of Brasilia, which was the first Brazilian University to adopt quotas for indigenous students under a Federal governmental social inclusion policy, was an ideal setting for indigenous students graduating from Intercultural Studies to be admitted to the Graduate Programme in Linguistics, specifically in the area of research named Theoretical and Descriptive Linguistics of Indigenous Languages. The greatest problem faced by this proposal was to convince faculty and students members of the Linguistics Department of the advantages of having indigenous students, since they are speakers of their own native languages.

But how would an indigenous student survive in an educational system requiring an "A" student profile? Opponents feared that the graduate programme

could suffer penalties, as the performances of such students did not correspond to the expectations of completed theses and dissertations. How could such students improve their linguistic knowledge if they are fluent neither in Portuguese nor in English, quite apart from their difficulties in dealing with the philosophical and epistemological scientific concepts they need to discuss in their scientific works?

All of these were difficult obstacles to face in trying to advocate for the admission of indigenous students into a Linguistics programme. The quotas for indigenous peoples were preserved despite strong negative reactions, including prejudice on the part of non-indigenous students worried about having their employment opportunities reduced as a consequence of the quotas reserved for indigenous students.

3.4   Five years of teaching indigenous students in Linguistics Graduate Studies at the University of Brasília

The outcome of the evaluation of the performance of indigenous students at the graduate program in Linguistics, viewed as a process, is highly positive, despite the difficulties they face. We can characterize the process as having six main phases: (1) facing the unknown, where students experience the torment of trying to acquire knowledge of linguistics; (2) starting to be familiar with linguistic analytic procedures and theoretical issues; (3) learning how to see and feel their native language from outside, as they say, (4) fascination at discovering what they could never imagine to know about the complexity of a language as a cultural tool, as conceived by Everett (2012); (5) reaching visibility in academia and bringing new knowledge to the academic setting, and (6) self-confidence as independent, indigenous researchers, together with consciousness of being far from the academically normative profile of an 'A+ white graduate student'; but also the object of anxiety generated by their incomparable knowledge of their own languages.

The five years of experience, up to the present moment, of indigenous students pursuing Linguistics studies at the University of Brasília provides important clues as to ways of improving Brazilian educational and linguistic policies for indigenous education. One of its main contributions is to underline the fundamental importance of training indigenous people as teachers and researchers at a high level of formal education. The indigenous students are very conscious that they are at the University to become linguists of their own languages, and then to go back to their communities to improve their school programmes, to develop more adequate native language teaching, and to initiate themselves the training of new educational providers.

Students' experience at the University of Brasilia, in both undergraduate and graduate studies, is making them increasingly aware of the endangerment risks affecting their languages and cultures. They are also increasingly aware of their political role as representatives of their communities, as well as representatives of indigenous people in general, in negotiations with the governmental sphere. They know how precious their linguistic knowledge of their own languages is for their struggle for self-determination. The indigenous students are also highly conscious of how much they still need to learn in the fields of Linguistics, Anthropology, and other disciplines.

4.   Conclusions

The recognition of the national linguistic diversity by the Brazilian Government is a prominent fact, and the indigenous educational policies associated with linguistic policies are nowadays realized through various educational programmes covering all levels of studies for indigenous people. There are also special programmes benefiting indigenous health, as well as indigenous culture, and all of them increasingly interact. All Ministries are responding, to a greater or lesser degree, to indigenous peoples' rights and needs.

A few years ago, the *Instituto do Patrimônio Histórico e Artístico Nacional* (IPHAN: Institute of the National, Historical and Artistic Heritage) started developing a project to establish a National Inventory of Linguistic Diversity furthering the recognition of Brazilian linguistic diversity. The idea is to promote each of the approximately 220 languages spoken in Brazil, including immigrant languages, sign languages, indigenous languages and Portuguese variants conceived as Afro-Brazilian languages, as a vital constituent of national heritage and patrimony.

As each indigenous language will be then recognized as heritage, it can be expected that there will be new Brazilian indigenous languages policies, and planning for their preservation. A major step already achieved by this project is the formal commitment of all Ministries to contribute to the development of these new policies. The IPHAN staff is also aware of the fundamental need for the indigenous communities' commitment. An important lesson already incorporated is that without community engagement, the programme will fail to attain its objectives.

In relation to language preservation and revitalization, the role of the indigenous linguistics teachers and researchers is fundamental, and new policies need to be developed in order to support the training of indigenous linguists at a high level of formal education.

All these policies and programmes have highlighted and recognised the multi-ethnic and multilingual nature of the indigenous population in Brazil, but more still needs to be done. The implementation of these policies must have a central focus on community engagement and empowerment of these groups. In this way it is possible to transmit a positive message to the new generation that is important to speak their languages. Linguistic diversity is a reality, and is now recognized not only for political reasons but also on the grounds of scientific and social rationale.

The approach we should advocate is that political agents, scientific researchers and teachers, and indigenous communities should work together in promoting not only the continuing documentation of indigenous languages, but also in promoting indigenous education and linguistic training as an additional way to strengthen indigenous language usage, and its transmission to future generations.

References

BRASIL. 1988. *Constituição da República Federativa do Brasil*. 5 December 1988. São Paulo: Atlas.

BRASIL/MEC/SEF. 1999. Ref. MEC/CNE. Resolução n. 03, 10 November 1999. <http://portal.mec.gov.br/cne/arquivos/pdf/CEB0399.pdf> (10 June 2010).

BRASIL/MEC/SEF. 1998. *Referencial curricular nacional para as escolas indígenas*. Brasília: MEC/SEF.

BRASIL/MEC/SEF. 2002. *Referenciais para a formação de professores indígenas/Secretaria de Educação Fundamental*. Brasília: MEC/SEF.

BRASIL. 1991. *Decreto Presidencial no. 26*.

BRASIL/MEC/SEF. 1993. *Diretrizes para a política nacional de educação escolar indígena*.

BRASIL/MEC. 1996. *Lei no. 9.394 – Lei de Diretrizes e Bases da Educação Nacional*.

BRASIL/MEC. 1998. *Referencial para a formação de professores indígenas*.

BRASIL/MEC/SEF. 2002. *Parâmetros em Ação para a educação escolar indígena*. Caderno de Apresentação.ed Luís Donizeti Benzi Grupioni .

BRASIL/MEC/SECADI. 2013. *Portaria no. 98, de 6 de dezembro de 2013*.

Cabral, A.S.A.C., Monte, N.L. & Monserrat, R.M.F. (eds). 1987. *Por uma educação indígena diferenciada*. Brasília: Fundação Pró-Memória.

Everett, D. 2012. *Language: The Cultural Tool*. London: Profile Books.

Gonçalves, M.F. 2013. Casa de escrever no papeo: a escola Tuparí da terra Indígena Rio Branco, Rondônia. MA dissertation, Universidade Federal de Rondônia.

Monserrat, R.M.F. 2000. Política e planejamento linguístico nas sociedades indígenas do Brasil hoje: O espaço e o futuro das línguas maternas. In *Questões de educação escolar indígena: Da formação do professor ao projeto de escola*, 1. J. Veiga & A. Salanova (eds), 127–171. Campinas SP/Brasília, DF: ALB/FUNAI.

Quintas, J. da S. (ed.). 1996. *O difícil espelho – Limites e possibilidades de uma experiência de cultura e educação*. Rio de Janeiro: Edições do Patrimônio.

Rodrigues, A.D. 2012. Contribuição para o Inventário Nacional da diversidade Linguística. Ms.

Rodrigues, A.D. 1985. The present state of the study of Brazilian Indian languages. In *South American Indian languages: Retrospect and Prospect*, H.M. Klein & L.R. Stark (eds), 405–439. Austin TX: University of Texas Press.

Rodrigues, A.D. 1986. *Línguas Brasileiras: Para o Conhecimento das Línguas ndígenas*. São Paulo: Edições Loyola.

Seki, L. 1999. A linguística indígena no Brasil. *D.E.L.T.A.* 15: 257–290.

Language rights in danger

Access to justice and linguistic (in)equality in multilingual judicial contexts

Liz Hales and Luna Filipović

University of Cambridge / University of East Anglia

This chapter provides a discussion of communication problems that arise in a multilingual legal context. We analyse witness interview reports and interviews from both the UK and the US in order to assess the difficulties that non-English speakers can face in an English-speaking justice system. The problems encountered indicate the need for the provision of adequate language support and improved professional training that will focus on particularly problematic lexical and grammatical contrasts for translation. We argue that people's language rights can be endangered as a result of the difficulties we discoverd, even for speakers of a major language (such as Spanish). We conclude that for the purpose of equality in access to justice these problems need to be addressed by both scientific and professional communities involved.

Keywords: access to justice, endangerment, English, interpreting, language rights, migrant, Spanish, translation

1. Introduction

This chapter explores the endangerment of the right to understand and be understood, for those whose first language is not Standard English, when they are processed through the criminal justice and immigration procedures in English-speaking countries. It explores the degree to which the impact of this is recognised in assessing the need for support and resultant service provision, to ensure that those whose cases are being processed have the same access to just outcomes as English speakers. This includes the provision of adequate professional interpreting and translating services and correct interpretation of statements for the officials working within these systems. Translation must take account of the impact of structural differences between language, as well as the challenges of interpreting

where intonation, gesture and facial expression[1] contribute to meaning-making in culturally specific ways. In terms of language rights in danger we include examples at different ends of the language spectrum. The cases discussed include minority languages, where the speaker may face the additional disadvantage of illiteracy, as well as "big" languages (i.e. those with multitudes of speakers around the world), like Spanish, who can also be disadvantaged when they find themselves in the communicative situation where English is the language of interaction.

This paper is based on original empirical data gathered in two English-speaking countries, with different legislative systems; the UK and the US. This comparative dimension allows us to highlight the extent to which the problems are pervasive, at different levels and in different justice systems. It also looks at the issues from the two different perspectives: that of a social anthropologist who has been a practitioner in the criminal justice system for many years in the UK (Hales), and that of a linguist whose work includes research in the criminal justice system in the US (Filipović). Both perspectives recognise the dynamics of the interaction between those in authority, the person who has been held in custody and the interpreter or translator, the impact of stress on the ability to think and speak clearly and the importance of the accuracy of the recorded statement in critical decision making.

In Section 2 we illustrate this in relation to the justice system in England and Wales with references to research findings by Hales and Gelsthorpe (2012). In Section 3 we provide some specific examples that cause problems in communication due to particular differences between two languages (English and Spanish) in the US justice system, based on previous research by the second author. In the final section we highlight the impact of service provision on outcomes, and propose remedies that ensure more effective communication between the individual and those working within the legal systems.

These areas of research should be seen in the context of other projects that tackle the relationship between language and law in relation to access to justice and infringement of language rights.[2] This area of work grows in importance with the impact of increased levels of international migration for work, study and to seek asylum.[3] The result of this, in a number of countries with a high GDP, is the

---

1. For example lack of eye contact, a sign of respect in many cultures could be wrongly construed as indicative of making false statements.

2. See Freeman and Smith (2013) *Law and Language. Current Legal Issues Volume 15* for a recent overview.

3. As exemplified in the annual reports made by the International Organisation for Migration (IOM)

tightening up of immigration control and growing numbers of migrants being held in prison and immigration removal centres, having been charged with offences in relation to their illegal entry, undocumented status, or illegal activities within their destination country. Unfortunately there is evidence that this growth in need has not been matched with resources to ensure that the language barrier is successfully overcome in the way these migrants are managed through the criminal justice and immigration procedures.

2. Facing the law without speaking the language: UK migrant prisoners' perspective

Hales and Gelsthorpe (2012) carried out an 18 month research project on the criminalisation of migrant women in England and Wales.[4] The goal of the project was to identify whether there are potential victims of trafficking, smuggling and work under duress in custody and, in the context of national data, give an indication of the extent of the problem. By identifying and monitoring these cases, the second task was to provide evidence on how such victims are managed within the Criminal Justice System and by the Immigration Criminal Casework Division and to what degree this is in compliance with the European Convention on Trafficking[5] and Human Rights legislation.[6]

From initial interviews with 103 migrant women held in prison and immigration removal centres, 58 women were identified by the researchers as potential victims of trafficking, work under slavery or servitude or abuse by smugglers and extensive data was gathered from this target group in terms of nationality, language skills, their socio-economic background and reasons for migration. Data was also gathered on their management through the criminal justice and immigration systems, with 73 follow up interviews, observation of 33 court hearings, examination of relevant paper documentation held by the interviewees, their legal representatives and others wherever this was feasible, and on-going communication by letter.

---

4. The term *Migrant Women* is used for those who have entered the UK to seek work or asylum, voluntarily or under duress.

5. Council of Europe Convention on Action against trafficking in Human Beings (May 2005). Since completion of this research this has been replaced by the EU Directive on preventing and combating trafficking in human beings and protecting its victims 2011/36.

6. For the full report on this research see Hales and Gelsthorpe (2012) *The Criminalisation of Migrant Women*.

A key finding of this research was the low level of recognition of victim status, at point of arrest and/or during the criminal justice proceedings, which should have led the courts to consider whether it was in the public interest to continue with the criminal charges. Of the 43 who were identified as victims of trafficking by the researchers, only 11 were processed through the National Referral Mechanism (NRM) to formally establish victim status and this did not happen for two of these women until their sentences were complete. Within this group, where action was taken, there were only eight positive decisions.[7] With women, who had worked under slavery or servitude or were transported by smugglers who withheld their travel documents, there were similar failures in recognition of the absence of knowledge or culpability in relation to the criminal act. During the period of the research there were comparable negative outcomes from asylum applications, with only two victims, both of whom had been formally assessed as children whilst in the adult estate, initially granted leave to remain within the UK.

These women and children[8] had complex and traumatic stories to tell and in seeking to identify reasons for low levels of victim recognition a number of factors were recognised; however, one of the recurrent themes was that of the language barrier. This together with low levels of education, lack of previous life experiences outside of their country of origin and ignorance of the foreign criminal justice and immigration systems led to further disempowerment. It impacted on their ability to understand, be understood and the development of trust which is critical to the management of victims of multiple trauma. Women's statements, written communications and court hearings observed by the researchers raised the following issues:

- Failure to recognise the need for interpreter and translation support, particularly where the arrestee appeared to speak some English.
- Recognition of need, but failure in provision.
- Exacerbation of problems by the use of technical English.
- Concerns around accuracy, completeness, impartiality, confidentiality and conflict of interest in relation to interpreter provision.
- Failures to ensure the work of interpreters was limited solely to language interpretation work.
- The impact of shame in disclosing experiences of sexual abuse through a male interpreter.

---

7. See National Crime Agency website for description of NRM procedures and conclusive grounds decisions.

8. From this point on we refer to the research sample as women, although it included two children.

- The lack of time to ensure accurate interpretation and translation of statements made in relation to the criminal charge and asylum applications.
- The impact of the physical court structure and proceedings in the level of comprehension.
- Important communications by letter to those in custody written in English.

These issues were exemplified at all the key stages of the criminal justice and immigration proceedings as outlined in the following research case material.

2.1   At point of arrest

Arrest was described by all those interviewed as the most critical and stressful time and those who spoke some English were often the most disadvantaged in relation to the language barrier, as disclosed in the following statements.

> "At the police station I was confused. They spoke quickly. They never asked if I needed an interpreter. I did not understand what was going on. I was crying … just wanted to tell them everything – I wanted them to listen and understand. If I had been able to talk then maybe I would not be here. The solicitor just said 'say no comment, no comment, no comment'. No one would let me talk to them."   (Yoruba speaker)

> "They were talking to me at police. I did not understand. Were all talking too fast they asked if I need an interpreter. I said Esan[9] – they said 'it is not on the system'."   (Esan speaker)

These statements were made by two victims of sex trafficking from Nigeria, both arrested for use of false documents. The second was formally identified as a child, nine months into her sentence in the adult estate. This was not recognised at point of arrest, nor was the fact that she was totally illiterate when asked to sign her witness statement.

Where interpreters were provided there were additional concerns with regard to their role and the impact of additional information they gave on management of the case. For example one Farsi speaker, who had been arrested for entering the country without legal documentation, despite the fact that she reported the theft of her passport by her smuggler as soon as she disembarked from the flight, stated that she heard the interpreter state in English that:

> "She says her English is not good – that in no true, all educated women like her speak and understand English."

---

9.   Esan is a tonal Edoid language of Nigeria.

In another case (a Columbian woman arrested for street robberies under the control of her trafficker) the same interpreter used at the police station was also used in court the following day, which is contrary to the national guidelines for use of interpreters in criminal justice matters.[10] The woman was upset by this and she clearly described the inadequacy of the treatment on that occasion:

> "I did not trust the interpreter used by the police and then by the court. In the police station she was not just interpreting – she was saying lots of other stuff to the police. She was used to interpret for all of us arrested and then my solicitor had to use her and then she was the court interpreter the next day."

There are further instances of unethical interpreter behaviour. Two women who spoke no English stated independently that at the police station the interpreter challenged them saying "Why are you lying?". They did not describe this question as interpreting what the police were asking, but rather as an independent question.[11] In another incident the defendant in court asked the interpreter to intercept on her behalf as an inaccurate statement was made to the court. The interpreter refused to do this saying it would be a "waste of time".[12]

2.2   Criminal court proceedings

One of the themes repeated in numerous interviews was the fact that the women interviewed did not understand what was happening in court, and this was made even more problematic where there was inadequate or no interpreting support when they were in dialogue with their legal representative. The unfamiliarity of the system exacerbated this problem as the following case shows.

---

10. Section 4.4.1 of the national guidelines state that; *It is important that so far as possible the interpreter arranged for court is not the one who interpreted at the police station either for the police or the defendant's solicitors at any stage prior to the court appearance. If however it is not possible to find another interpreter (where for example the language is rare) then the court and all parties must be notified of the intention to use the same interpreter for the court proceedings and agree to that course of action.* In this case the language was Spanish and therefore access to alternative interpreter support should not have presented a problem.

11. If this question *why are you lying* had in fact been made by the arresting officers, this would indicate the type of confusion as to what was being interpreted and what was being stated independently by the interpreter, as evidenced in the research material from the US (Section 3).

12. This statement error was noted and the interaction was observed by the researchers in the court room.

> "I spoke to the solicitor on video link before going to court. I said I do not understand and want a Dutch speaking solicitor."

She restated this in her request to change her legal representative and the following is an extract from the letter sent in response:

> "You should write to the court again, giving them a much stronger reason why you want legal aid changed. … Nobody in this firm speaks Dutch."

Her request to have her legal aid transferred, which she had to make in English, a foreign language, was never taken as an indication that interpreting support was essential. When observed in court there was no interpreter present and she was not asked formally as to whether she understood the proceedings.

When interpreters are used in court, the interpreter has to make an oath to the court. In none of the observed criminal court proceedings cases, was the oath made by the interpreter and then repeated in the language of the defendant. In only one case (a video link court hearing) did the court formally identify the role of the court interpreter to the defendant. In a second case the interpreter stopped proceedings to remind the lay bench that this should happen.

Where interpreting support was used, there were two types of interpreting provision observed. The first was a short consecutive interpretation where the bench, that is the judge or magistrate, was involved in direct dialogue with the defendant. This was done to confirm name and date of birth of the defendant, outline the charges, advise on the date of next hearing, to hear the plea and, in the context of Crown Court trials, where the defendant was being cross examined. The second type was simultaneous interpretation, where the court expected that the matters discussed would be interpreted for the defendant so that she was aware of what was happening.

Failings in both types of provision were observed and also identified by the women themselves. In two hearings that were observed the interpreter failed to appear. The court recognised the need, as neither of the defendants spoke or understood any English. However as proceedings in court that day were arrangements for a further adjournment, and there was to be no direct dialogue with the defendant, the courts decided to *proceed anyway*. In this case there was no interpreter support for a de-briefing after the court appearance. In other cases the women reported that by the time they met with their solicitor in the court cells after the hearing the court interpreter had left the building and they did not understand the outcome. Observations of meetings between women and their legal representatives confirmed this failure and the fact that the use of incomprehensible legal jargon exacerbated this problem.

In all cases observed, where interpreter support was needed for co-defendants speaking the same language, this was provided by the same court interpreter. In only one hearing was the judge observed requesting that the Crown Prosecution Service take account of the need for adequate interpretation in terms of how they proceeded, and no cases were observed where the defendant felt they could stop proceedings for clarification. A common statement by those interviewed was that they did not feel they were benefitting from a full interpretation of all that was being said, and they tried desperately hard to follow the gist of the proceedings as well as listen to the words of the interpreter.

This ability to hear, understand and engage in dialogue with the legal representative through the interpreter was also impacted on by the physical court structure and layout.[13] The defendant was normally positioned at least two metres behind her legal representative and thus unable to get his/her attention. In all but two criminal cases the defendant was held in a dock at the back of the court, behind a ceiling-high glass screen, where she was joined by the interpreter. Two court interpreters independently talked with the researchers about the impact of working within this screened off area, stating that they also faced difficulty in hearing all that was said as "the speaker system is often faulty". Another pointed out that "if the officer in the dock jangles his keys it is so easy to miss key words".

Of equal concern is the fact that there was no evidence of third party checks that the interpretations were full and accurate. Three interviewees, who were bilingual, stated that there were many mistakes in the interpretations they heard in court provided by court interpreters for co-defendants. They expressed frustration that they could not raise the attention of the court in relation to this.

In addition, those most in need of interpreting support, such as those from Vietnam, were often arrested with others who had control over them, or whom they could not trust in terms of passing information back to those who had trafficked or worked them and to whom they were in debt bondage. They were therefore fearful of making any statements that, if not dealt with in confidence, could prejudice their future safety. The fact that the same interpreter was used by the different legal representatives for the co-defendants, for pre-trial briefings at the court, exacerbated these anxieties.

Talking about these procedures, views routinely expressed by those in custody were confusion, frustration and disempowerment. A common statement to the researchers was the feeling that they were effectively "off stage" and not involved or able to impact on what happened within the court. This is despite the fact that any person charged with a criminal offence in the UK has the right to "have the free assistance of an interpreter if he cannot understand or speak the

---

13. For analysis of the impact of this see Mulcahy (2011).

language used in the court."[14] It is therefore the duty of the Bench[15] to ensure that the defendant understands what is being said and if they become aware of the need for interpreting support to ensure the case is adjourned for this to be arranged. However there were no routine checks observed in the hearings observed.

The impact of the language barrier was recognised in one trial observed where the judge advised the jury that:

> "You must make allowance for the fact that the defendant has given evidence through an interpreter. It is more difficult to get across the real flavour of what you want to say when it is done through an interpreter. Someone … I cannot remember who it is … once said that you lose the poetry, that is the feeling, the nuances, the clarity of what a witness said when his or her evidence is interrupted. So you must make allowances for that as you think appropriate."

At this trial the level of interpreting was also very high, with the interpreter explaining at times that he needed to explain the term used by the court, and asking at one stage, before deciding on the intonation he should use, whether the statement made in cross examination that he was interpreting had been posed as a question.

Unfortunately however the above case was the exception and there were a worrying number of statements outlining lack of effective interpreting support. Even where the interviewee spoke English as a second language and did not require interpreting support in the context of the research interviews, key legal terms, sometimes including the actual meanings of the charges, were not understood. This is exemplified in the following two statements:

> "Although I understand English there are professional words I do not understand. I did not know what is meant by 'conspiracy'" (to which she was advised to plead guilty).

> "I wanted to say to the judge, I do not understand the English you are talking. I do not know what is going on."

It is also of significance that in three of the cases studied in this research the original conviction and sentence was finally overturned. However in none of these Appeals was lack of adequate interpreter support listed in terms of procedural failings by the Appellant or formally recognised by the Court of Appeal.[16]

---

14. In compliance with Article 6 of the European Court of Human Rights

15. The Bench is a legal term used to describe the Judge or Magistrates hearing the case.

16. It is of relevance that this factor was not recognised in other appeal hearings observed by Hales, following the publication of the 2012 research.

2.3   Immigration proceedings

In asylum procedures, the lack of translation also disempowered the asylum ap-plicant's ability to ensure that she had disclosed all the relevant information, or to challenge the grounds for negative decisions. For example after the full asylum interview, carried out with the support on an interpreter, one woman stated:

> "Yes they showed me the interview notes – but it was impossible to check if they were right. They were all written in English and I could not read some of the handwriting."

Several weeks after this interview this woman received an asylum decision which was one of refusal. This standard refusal letter, written in English, stated that the applicant had five working days in which to appeal and if she wished to appeal she had to complete the attached document. This was a lengthy document asking for details of the basis of the appeal and contact details for the immigration solicitor. It was not an easy document to understand for an English speaker. In relation to these procedures it is of relevance to note that Article 10 (1) of the Asylum Proce-dures Directive (APD) which requires:

> "Member states to inform asylum applicants of the decision on their appli-cation in a language that they must reasonable be supposed to understand, unless they are represented by a legal adviser or free legal assistance is avail-able, in which case the requirement may be waived."[17]

As with many of the other asylum seekers in custody she faced huge challenges in accessing legal support and her limited English further inhibited her ability to access her representative by phone or letter and thus seek explanation of the latest communication from Immigration.

Some of the same issues that were apparent in criminal court hearings were also observed at Immigration and Asylum Tribunals.[18] In one of the cases the level of interpreting was so poor in terms of accuracy and failure to replicate the intonation of the answers that the applicant asked to be able to speak directly to the judge in her limited English. This was denied by the judge who also seemed

---

17. EU member stated remain committed to the APD under Article 68 of the Treaty and Func-tioning of the EU.

18. These were the next stage of appeal for those refused asylum. For those in custody at the time of these hearings, they were held in a closed dock area of the court and they were not allowed to sit beside their legal representative in the open court, as normal applicants were. It also meant that in some cases they were physically separated from the court interpreter.

unaware of appropriate protocol where interpreters are used, in terms of ensuring that face to face contact is between the applicant and the court, when she stated:

> "We have an interpreter in court. I will ask my questions to the interpreter who will interpret them to you. You will answer to the interpreter who will interpret to me what you have said."

In the context of the letters from the Immigration Department, which stated that asylum had been refused and which were seen by the researchers, one of the key reasons stated was often "lack of consistency" between statements made at initial immigration interviews and full asylum interviews. Linked to this is the fact that a dominant reason for asylum applications within this group was the experience of rape, and the fear of future rape and sexual abuse. Women interviewed talked about *shame* and *fear* involved in disclosing these issues in front of a male interpreter at the initial interview.[19] In relation to the following extract from a letter from Immigration it is also possible to conclude that these apparent inconsistencies may have more to do with the actual process of interpretation.

> "In your account of the rape you first stated 'when we got somewhere they took me. They took me from a different place they took him. From there they raped me'. You also state, 'after they raped me they took him to a different place' … it is considered that you have supplied conflicting accounts of the rape and consequently it is difficult to accept that this is a true account of a real event."

One explanation of this is the fact that the victim was highly traumatised, with resultant difficulties in narrating all that she had experiences in a logical coherent manner. However it is also possible that, in the process of interpreting, some personal pronouns or prepositions have been confused (such as they "took me from a different place they took him" vs. "they took him to a different place"). In any case, the excerpts "they took me from a different place they took him" and "from there they raped me" are inadequately rendered into English and it is this inadequate interpreted English that appears to be the major source of confusion. However in this case as with others in the UK there was no routine practice of comparison between the statement in the original language and the interpreted statements, nor the production of bilingual transcripts that could be checked by a control translator, as is practices in the US for police interviews.

---

19. For fuller discussion on the issue of inconsistency and gender specific persecution see Quernton (2012, Section 3).

3.   Police interviews in the US: Barrier of translation

The importance of the translation and interpretation quality *per se* is an issue that is of crucial relevance in the context of police interviews (see Drugan 2013 for a general discussion on ethics and translation quality). This is particularly evident in the research results of Filipović (2007, 2013a), who investigated the language provision practices in the judicial system of the state of California, with particular focus on non-English speaking migrants in the US courts of law. The United States, like the UK, are under significantly high pressure due to immigration. In particular, some areas of the US have specific immigrant groups that are present in large numbers. Such is the presence of Spanish-speaking citizens from Latin American countries in the state of California, where Spanish-speaking interpreters for both police and court interviews are in constant, relentless demand. Filipović (2007, 2013a) reports on the fieldwork findings from a number of jurisdictions within California (San Francisco, San Jose and Oakland), where over 17 000 pages of police (127 files) and court interview (69 files) transcripts were collected. The police interviews and the resulting transcripts were bilingual (Spanish original with English translation). The court interviews were carried out bilingually with an interpreter, yet however, only the English translation is recorded there. Police transcripts in the US are regularly produced as bilingual documents verbatim for the interviews with non-English speaking subjects, while the court transcripts are produced as monolingual (English only) files even when the defendants and witnesses speak languages other than English. Crucially, police interview transcripts are additionally checked and translated post-interview by an independent translator (different from the one present in the original interview). In this way, the quality control of police interview transcripts is impressive and the quality of the service is very high. This can serve as an example of good practice for others to follow. The control translator has the opportunity to correct interpreting errors or clarify the use of certain words or expressions.

Unfortunately, the same procedure is not available for court proceedings due to the incurrence of high costs, especially with lengthy criminal trials. The problems in access to justice that arise as a result of having monolingual records of multilingual communication have been widely documented. A number of previous studies have addressed many relevant issues that arise in a multilingual courtroom (e.g. Berk-Seligson 1990), or other instances of multilingual interviewing (e.g. police investigation (Berk-Seligson 2009) and community interpreting (Hale 2004)). Numerous researchers have tackled the general problem in multilingual legal cases that stems from the extremely difficult role that interpreters on occasion have to play. For instance, Berk-Seligson (1990) lists a number of features that characterize court interpreting (e.g. hedges, insertions, hesitations, etc.). These

features underlie the perception of witness testimony style as either powerful or powerless. A further example of cross-linguistic contrasts in translation is Hale's study (Hale 2004), which documented the difficulty of translating tag questions from English into Spanish.[20] On this occasion and for the purpose of this section of the paper, we focus on the police data since it is only possible to discuss issues of language contrasts and difficulties in translation based on bilingual transcripts, which are the police interview ones.

We draw attention to two issues relevant for professional practice that were evident in the US data and that bear consequences for the endangerment of language rights: (a) police officers acting as interpreters and (b) points of contrasts between languages that create particular difficulties even when the interpreting is carried out professionally.

3.1   Police officers as interpreters

Investigative police interviews in a bilingual communication setting are effectively more challenging than monolingual ones. The problems that arise as a result of having bilingual police officers as interpreters instead of using professionals has been documented, acknowledged and explicitly addressed by the US judiciary (as well as some other jurisdictions, e.g. the UK and Australia; see Abad Vergara and Filipović (submitted)), and this judicial awareness can serve as an example for other justice systems throughout the world. Such high standards of service provision are not easy to achieve, or always adhere to, but an exemplification of problems that are caused by having officers perform dual roles (that of an investigator and an interpreter) can lend support to the argument against such practice.

Abad Vergara and Filipović (submitted) carried out a case study analysis using an authentic case of a suspect being interviewed by two police officers, one of which was introduced as an interpreter. The police interview was carried out with a Spanish-speaking suspect charged with a serious offence (sexual assault), and one of the two officers was also acting as an interpreter throughout the interview. During the interview, the suspect was increasingly confused by the impossibility to understand which questions were coming from the principal investigating officer, which ones were coming from the officer-interpreter. The interview in question took place in the United States in 2000; that is, before the introduction of new regulations regarding the employment of professional interpreters in police interviews. The US Department of Justice had researched the employment

---

20. The most recent holistic overview of themes and methodologies in the field is given in Gibbons (2011).

of language services and provisions in legal cases and as a result they "elaborated a set of prohibitions on the use of non-professional interpreters by law enforcement agencies" in particular, the use of bilingual officers or employees regardless of level of bilingual competency (Berk-Seligson 2009: 13). Abad Vergara and Filipović draw attention to the ways in which both linguistic accuracies and the lack of impartiality on the part of the interpreting officer could contribute to blame attribution and lack of neutrality required for the interpreter as a result of the interpreter's dual role (i.e. interpreter and police officer).

For instance, there are constant references to the suspect in the third person, whereas professional interpreters are required to use the first person when speaking in the words of the original speaker:

(1) a.   Suspect: Oh, no le pregunté.
           Translation: Oh I didn't ask her.
     b.   Police officer-interpreter: He never asked her.

Similarly, subtle differences in the choice of words by a person whose proficiency in the foreign language is not attested can cause differences in the interpretation of the statements in the original and the translation, as in:

(2) a.   Suspect: […] ya de ahí pues ella quería ir a la escuela.
           Translation: […] from there well she wanted to go to the school.
     b.   Police officer-interpreter: […] she had to go back and I took her back to school.

The difference between the modal meaning in Spanish and its translation into English adds an extra layer of meaning, whereby the police officer-interpreter conveyed the message as an obligation (i.e. *had to go*) whilst the suspect speaks of a desire (i.e. *wanted to go*). The suspect is saying that the alleged victim wanted to do something and, as evidenced later in the script, he complied with her wish and walked her back to school. The translated statement, on the other hand, indicates that the alleged victim had the obligation to go to school but does not indicate that it was actually her wish that the suspect complied with. There are further examples that illustrate inadequate level of proficiency and skill of the police officer-interpreter, e.g. in his use of verbs and constructions as evident in the following example:

(3) Police officer-intepreter: *Cuando usted la caminó para la escuela…
     Translation: When you walked her to school…

This construction of using motion verbs (such as *walk*) transitively (*to walk somebody somewhere*) does not exist in Spanish. It is a calqued English construction,

and as it stands in Spanish, it makes no sense and, furthermore, causes confusion. The police officer-interpreter uses this construction on a number of occasions and the suspect hesitates and falters each time, asking for clarification on occasion (i.e. by saying "Comó?" = "What?" or similar). We know that pauses and hesitations have negative consequences on communication in general (Dingemanse & Enfield 2014; Roberts, Margutti & Takano 2011), especially in judicial contexts where it creates the impression of a speaker being powerless and consequently less convincing or trust-worthy (Berk-Seligson 1990). The cause of such negative judgment can create a language barrier, as exemplified here.

The crucial insight that stems from the results of the Abad Vergara and Filipović study is that it is not enough to be a bilingual speaker in legal interviews and that specialist training, especially in the field of legal translation, is extremely beneficial (see also Abad Vergara 2014 for a detailed study on the necessity and benefits of professional training). Non-native speakers are put in a disadvantageous position from the very start, in a first interview with law enforcement and then further throughout the judicial process in courts, where the original statements they make are never recorded. The excellent practice of making bilingual transcripts in the United States police interview contexts is extremely helpful for the purpose of revealing the kind of disadvantage that non-native speakers may face. Further issues stem from the fact that languages differ with respect to ease or difficulty with which their speakers can express certain meanings. We turn to those language-specific difficulties and the barrier they create in translation-assisted legal communication.

3.2   Language barrier due to language contrasts

Language contrasts present communication participants with additional difficulties with regard to the adequacy with which original information is rendered in translation. For instance, two typological dimensions along which English and Spanish are very different are the use of verbs and constructions for the expression of motion and causation. These differences have been widely documented by numerous scholars (e.g. Slobin 1997, 2000, 2004, 2006; Filipović 2007, 2013a, 2013b; Gibbons 2003). In this section we present examples that illustrate the typological contrasts that are responsible for the inaccuracies in translation and that are not necessarily due to interpreting incompetency but rather to the different habitual practices in each language that are conditioned by the specificities of grammar and use in each language.

3.2.1   *Motion expression in legal translation*

Spanish and English offer different strategies for speakers to use when verbalizing motion events. In English, speakers habitually say where and how something or someone moved (e.g. *Harry skipped into the shop*) while Spanish speakers tend to say only *where* something or someone moved, omitting the information about the *how* (i.e. the manner of motion; as in *Javier entró en la tienda (brincando)* = *Javier entered the shop (skipping)*). This is due to the fact that the manner of motion is expressed in an obligatory element in the English sentence (the verb) while the Spanish verb is used to express the path. The information about the manner, if at all provided, is given in a non-obligatory element, the gerund (*brincando-*skipping) and can (and often is) excluded from the expression of a motion event in Spanish (see Slobin 1996, 2006). This typological difference has numerous consequences for the description of motion in the two languages. Namely, Spanish translators often omit information about the manner from English: English translators add manner information even though it is absent from the Spanish original because their pattern of expression involves manner verbs and they feel obliged to provide manner information. A constant use of path verbs, as is done in Spanish, would sound somewhat unusual in English (e.g. *He ascended the stairs running* instead of *He ran up the stairs*). These contrasts have been documented in extensive research data from a variety of contexts of use (literary texts, spoken discourse, metaphorical language; see Slobin 2006) and at different stages of language acquisition (infants, children, adults; see Filipović & Ibarretxe-Antuñano 2015 for a detailed overview). Filipović (2007, 2013a) were the first studies that document the effect these specific typological differences on the understanding of communicated and translated message in a legal context.

For example, even in cases of sexual harassment and knife attack, which are events that must have occurred with highly salient (probably intense and aggressive) manner of movement, information on manner is conspicuously absent throughout the description of events:

(4)   Me salí de la oficina y me fui. Y él se fue atrás de mi, se fue, pero él se fue para allá y yo me vine para allá
Translation: I got out of the office and I left. And he went after me, he left, but he went over there and I came over there.

(5)   … y yo le caí atrás, lo vi que traía la, la, la navaja y le caí atrás y cuando le caí atrás, muchos le caímos atrás para agarrar al que agredió el muchacho.
Translation: … and I took off after him, I saw that he had the, the, the knife and I took off after him and when I went after him a lot of us went after him to grab the guy who had attacked the guy …

The descriptions of the two situations in example (4) and (5) also seem to lack dynamicity and intensity, which undoubtedly accompanied the reported events that relate to a sexual assault. However, the victim in that case was using the usual Spanish pattern of expression, which does not involve manner detail, potentially crucial for the creation of the impression regarding the speed and flow of events that could explain the situation better, for example, why the victim was unable to flee or escape her assailant. This lack of detail may have an impact on the victim's case, but the victim's language does not encourage the provision of such information habitually. Therefore, such information, about the manner in which the events unfolded, has to be sought explicitly.

While these accounts sound natural in Spanish, the absence of manner verbs makes the account sound very untypical in English. This is why in the process of interpreting the information about the manner of motion can be, and often is, spontaneously added, because it is the most natural way to lexicalize motion events in English, as we can see from the following example:

(6) Original: pero … salió por la puerta detrás.
     Literal translation: but … he exited through the back door.
     Official translation: but he … ran out via the back door.

The dynamicity of the situation from which the example (6) was taken induced the interpreter to add manner to describe a chase of the suspect even though the witness did not mention it explicitly in Spanish. The consequence of this is the possibility to impede the identification of a suspect or his whereabouts since the suspect could have run for a while but exited via the back door walking in order to avoid suspicion. In this specific case of a chase in the streets no information was made available in the Spanish original witness description on how quickly the suspect or the witnesses that followed him were moving. As a result, the situation can be interpreted as more dynamic in the English translation.

Information about the manner of motion is very important because it allows us to speculate about the suspect's physical state and location (e.g. if the suspect was running all the time, he could be tired and hiding in the search area; he would have gone further from the crime scene if he had run than if he had limped; if he had run, it means he had not been wounded or hurt, etc.). The communicative consequence is that we could draw different conclusions about a described event from the Spanish original and its English translation respectively. In the case of pattern-clashing such as this between English and Spanish, it may be useful for the interviewers to explicitly encourage speakers of languages like Spanish to provide information about the manner during interrogation.

3.2.2   *Causation expressions in legal translation*

Another relevant dimension of language contrasts is that of the expression of causation. The difference in how causation can be expressed in English and Spanish has presented interpreters with further significant difficulties that are hard to overcome in a high-pressure situation such as police interviews of suspects. One particular construction has been highlighted by Filipović (2007, 2013b), since it best illustrates the profound impact on the case outcome that this typological dimension may have. English is typologically an agentive language, where agents are clearly marked by their syntactic position in the subject slot. English is a language that clearly expresses agents, but does not oblige its speakers to state clearly whether the agents performed the action voluntarily or non-voluntarily (e.g. *Bill dropped the bucket*). Spanish on the other hand has two distinct constructions (discussed below) that clearly indicate whether the action was performed with intention (voluntarily) or without intention (non-voluntarily), but agents in that language are not obligatorily expressed.

This typological difference was shown to have profound effect on witness memory in experimental psycholinguistic studies. For instance, Fausey and Boroditsky (2011) have shown that English speakers remember agents better in both voluntary and involuntary causation events, and they had also confirmed (Fausey & Boroditsky 2010) that the explicit causation expressions (such as 'X broke Y') elicit more direct blame implication than the non-causational expressions used to describe the same events (such as 'Y broke'). Filipović (2013b) has shown that Spanish speakers remember better the difference between which actions were voluntary and which were not, in line with the Spanish speakers' language preferences of distinguishing clearly between the two action types. Namely, when the action is performed on purpose both languages have similar options as seen below:

(7)   a.   John threw the bottle.
      b.   Juan botó la botella.
           Juan threw the bottle.

However, when the action was performed *accidentally*, English and Spanish show an important difference in expression, whereby Spanish has a more precise construction for which English does not have an adequate translation equivalent.

(8)   a.   Se le cayó la botella (a Juan).
      b.   To-Juan-it-so-happened-that the bottle fell. (i.e. Juan did not do it on purpose)
           Juan dropped the bottle (non-intentionally).

In a case of police interviews, one particular case can be used to illustrate the importance of drawing attention to this typological difference and the need for its precise translation (see Filipović 2007 for details). 'Se me cayó en las escaleras' (meaning 'to-me-it-happened-that-she-fell on the stairs') was used extensively by a suspect who was describing what had happened to the victim and it was repeatedly translated as "I dropped her on the stairs" in English. The English expression *I dropped her* could refer to both intentional and non-intentional dropping, it is ambiguous and can therefore legitimately be used as an equivalent for the translation of the non-intentional expression in Spanish, so the interpreter was not at fault here strictly speaking. Throughout this particular interview the suspect was using the non-agentive expression in Spanish ("Se me cayó" = "It happened to me that she fell" or "I dropped her accidentally") that clearly signals lack of intentionality on his part (as in the example (8)). In the English translation however this clear non-intentional meaning was not available. In addition, the police officer did not make the questioning easier by asking "Did she fall or did you drop her?" It was obvious in the analysis that the interviewing officer used the verb "fall" in English for the non-intentional action and the verb "drop" for the intentional act. However, to the question "Did she fall?" the suspect responded "No" because technically, the victim did not fall, she was not in control of the motion. Crucially, the suspect responded "Yes" to the second part of the question, to the interviewer's intentional "did you drop her?" but his response was, in fact, the non-intentional, accidental version of "I dropped her" ("Se me cayó"). This was not understood by the police interviewer and the assumption was that the suspect was admitting guilt for an intentional act of throwing his victim down the stairs. In some US states, like California, which still administer the death penalty, admitting to an intentional act of murder is not something that can be treated lightly and the imprecise translation that leads to the understanding that suspect was doing precisely that is potentially highly detrimental, not just to the case, but also to the suspect's life. This is not to say that the interpreter was necessarily doing a bad job since, as we explained, the English phrase "I dropped her" can refer to both intentional and non-intentional act of dropping. Rather, due to the typological difference between English and Spanish in this domain, it is possible to leave certain important information as ambiguous and liable for wrong, and potentially extremely harmful, interpretation. Neither the interpreter nor the police interviewer was aware of this problem during the interview and it is not our aim to lay blame on any interview participant here. This communicative context is particularly sensitive and highly stressful, this must be borne in mind. Our goal here is to highlight the key point, which is the need for more empirical studies in this vein in order to detect what specific language contrasts create serious difficulties in translation between any two languages, resulting in serious misunderstandings.

What this example shows clearly is the central importance of raising awareness about language contrasts for the understanding of what a suspect is or is not confessing to. Furthermore, studying patterns in language use, characteristic of two different language types, and the speakers' habitual preference for certain constructions, can help us detect the exact points at which problems in translation may occur between any two languages. This is achieved by studying both the finely grained typological differences between languages and the analysis of real, authentic interactions among all participants in a social context (e.g. police interviewers, witnesses and interpreters) as we illustrated in this paper. In this way we can improve the focus of professionals involved in the process of collecting information that may be used as evidence.

One example of good practice in the US context is the employment of *control translators*, especially in cases of serious crimes. The role of a control translator is to check the transcripts and make sure that the original interpreting that took place was correct, accurate and faithful to the original statement of a witness or a suspect. The importance and value of this procedure becomes clearer and more convincing if we look at the following example:

(9)   Original: De ahí *la llevé* para la calle arrastrando y la puse sobre las, sobre las gradas y ahi *se volvió a dar el otro* golpe así, pero ya estaba muerta.
Transcript translation: Okay, and then he *grabbed* her and he try to take her out of the house and *put* her on the sidewalk.
Control translation: From there I *took her* to the street, *dragging her* and I put her on the, on the stairs and *she hit herself like this again*, but she was already dead.

The control translation highlights the differences between the original text and the interpreting registered in the transcript that took place before the control translation. We get different imagery based on the two translations, the latter of which, the control one, is the correct one, while the interpreted text that was given during the interview contains numerous lexical and grammatical errors. Namely, the words such as "grabbed" in the interpreted statement instead of "take" that the witness used can create an impression that the action may have been more violent than originally described. We know from previous research (Trujillo 2003; Ibarretxe-Antuñano & Filipović 2013) that the use of more complex and intense manner verbs can affect jury opinions about the events described. Moreover, the information about how the victim sustained further injuries (control translation: "she hit herself like this again") is missing from the original interpretation into English.

Packaging information in a language-specific way is so deeply rooted in our everyday experience and interaction with the world around us that we are often

unaware that we are doing it, namely organizing information according to a certain entrenched underlying system of words and rules. Specialist training targeting specific points of serious conflict between two languages can target this problem and prevent it from occurring in the future. In addition, if we are carrying out the extremely stressful job of interviewing and interpreting, we are naturally inclined to revert to the comfort of our typical and familiar linguistic frames. By this we mean that, when under pressure, people in general rely on entrenched stereotypes, including the linguistic ones (see Mendoza-Denton 2010). As was argued in this section, the typical way of describing the situation of dropping somebody that was carried in English is indeed the construction 'I dropped her' and the interpreter was not really completely wrong in the choice of expression when he used that ambiguous construction to render the non-agentive Spanish expression 'to-me-it-happened that she fell'. However, the potentially very damaging ambiguity that remained should have been clarified. An additional problem for interpreters is the uncertainty in their position, since they are required to refrain from any intervention, and the extent to which they are allowed to stop, or detract from the flow of, the interview proceedings can vary depending on the legal counsel in charge. The need to clearly explain the realm of intervention that the interpreter can have in advance would be immensely helpful to those professionals, and it would facilitate their hard job and add further sense of value to their service, in addition to a better quality of the translation itself. Moreover, it would reduce the gap between the service available to native speakers and those disadvantaged categories of non-native speakers discussed in this chapter.

4.   Conclusion

In Sections 2 and 3 of this article we have looked at access to justice for those whose first language is not English in two very different settings; however there are common themes in relation to this. These include the need for adequate interpreting support from point of arrest through to the conclusion of the court procedures, to ensure just outcomes. This would appear to be most significant in relation to establishing not simply whether a criminal act has been committed by the person charged, but rather the role that the defendant has played in this action including awareness and intent. For example, with those trafficked or smuggled into the UK, the use of a false document is easily established. However, what is not so easily established, and what has the greatest impact on outcome, is intent and/or awareness of the fraudulent nature of the documentation and the level of duress. This is even more critical when we look at the impact of intent on those charged with murder in California.

A second theme is ensuring that service provision is at an acceptable level, that there is very clear boundary management of the role of the interpreter and that legal professionals have adequate training in their use. A third is the recognition that communication is a two way process and as well as ensuring the foreign language speaker understands the process, it is critical that those investigating and prosecuting cases are correctly advised in terms of statements made, which necessitates interpreting activity which takes into account the impact of language structure and intonation. There is evidence in the US that this is taken very seriously in relation to bilingual documentation and third party checks of police interviews. However, these precautions are lacking in other areas of work.

The relevance of the findings is enhanced if we look at current data which indicates the level of need. The latest Ministry of Justice data for England and Wales shows that foreign nationals represent just under 12% of the total prison population.[21] There is no recorded analysis of literacy and levels of spoken English within this group, but as pointed out in Hales (2014), English is not the mother tongue in the top ten nationalities listed. Of equal significance are the numbers of Gypsy, Roma and Travellers in custody in the UK, where high levels of illiteracy exacerbate the language barrier.[22] Within the US the percentage of foreign nationals is slightly higher, at 14.6%.[23] However published data would appear to indicate a dominance of Spanish speakers within this group with almost 17% of the total population being those of Mexican nationality.

Within both countries there is formal acceptance of the rights of non-English speakers, but this paper attempts to demonstrate that there are still many ways in which those who are not fluent in English feel that they do not have the same access to what they view as just outcomes. This may be a result of numerous factors including their perception of the impact of their status as a migrant or outsider, the unfamiliarity of the judicial system and the additional impact of court decisions on future residential status. However in this paper we propose that the key factor is that of language. The examples we discussed come from English and Spanish, but we are confident that our study approach could be applied to other languages, whereby the relevant typological language contrasts and their manifestations and consequences in discourse and translation can be highlighted for any other language pairs.

To remedy the impact of this, good interpreting and translating facilities as well as additional time to process cases is critical. However, to keep abreast of a

---

21. Data from Ministry of Justice (April 2016).

22. For fuller analysis of this see Hales (2014) *The Language Barrier to Rehabilitation.*

23. Data from U.S. Federal Bureau of Prisons *Inmate Citizenship 30th August 2014.*

growing demand means additional funding and in the UK attempts were made to cut the cost of interpreter fees though provision of a single contract in 2011. Unfortunately this was recognised as having negative consequences including costly delays in processing court cases where there were failures in interpreter delivery.[24] Where resources are undoubtedly limited the response must therefore be targeting areas where the consequences of poor translation are most severe. Of equal importance is raising the awareness of those who manage these cases of the impact of failures in delivery.

Acknowledging that the danger of injustice exists is the first step forward, and this paper is hopefully making this kind of step. In addition, the real life examples of good and bad practice, as outlined in this research material, can be the key to work-focussed training for those working in these fields. However, there is still a huge need for further research that engages with the stages in criminal and immigration proceedings, whose outcomes have major impacts on the speakers of other languages. Communication and mutual understanding is key to all these procedures and our goal should be that of a fair playing field, where participants do not feel that they are in danger of not understanding and of not being understood. We hope that this and future research projects will help remove the language barrier on the road to equality in access to justice.

References

Abad Vergara, S. & Filipović, L. Submitted. Separating investigation from interpretation: The issue of mixing the interviewer and interpreter role in police interviews.

Abad Vergara, S. 2014. A Study of Interpreter Training in the UK on a Legal Context: Issues of Language Contrasts, Policies and Practices. MA dissertation, University of East Anglia.

Berk-Seligson, S. 1990. *The Bilingual Courtroom: Court Interpreters in Judicial Process*. Chicago IL: University of Chicago Press.

Berk-Seligson, S. 2009. *Coerced Confessions – The Discourse of Bilingual Police Interrogations*. Berlin: Mouton de Gruyter. doi:10.1515/9783110213492

Council of Europe Convention on Action against trafficking in Human Beings. May 2005. <http://www.coe.int/trafficking>

Dingemanse, M. & Enfield, N.J. 2014. Let's talk: Special report. *Scientific American Mind*, September/October: 64–69. doi:10.1038/scientificamericanmind0914-64

Drugan, J. 2013. *Quality in Professional Translation*. London: Bloomsbury.

Fausey, C.M. & Boroditsky, L. 2010. Subtle linguistic cues influence perceived blame and financial liability. *Psychonomic Bulletin and Review* 17(5): 644–650. doi:10.3758/PBR.17.5.644

Fausey, C.M. & Boroditsky, L. 2011. Who dunnit? Cross-linguistic differences in eye-witness memory. *Psychonomic Bulletin and Review* 18(1): 150–157. doi:10.3758/s13423-010-0021-5

---

24. See House of Commons Justice Committee 2013.

Filipović, L. 2007. Language as a witness: Insights from cognitive linguistics. *Speech, Language and the Law* 14(2): 245–267.

Filipović, L. 2013a. The role of language in legal contexts: A forensic cross-linguistic viewpoint. In Freeman & Smith (eds), 328–343. Oxford: OUP.

Filipović, L. 2013b. Constructing causation in language and memory: Implications for access to justice in multilingual interactions. *International Journal of Speech, Language and the Law* 20(1): 1–19. doi:10.1558/ijsll.v20i1.1

Filipović, L. & Ibarretxe-Antuñano, I. 2015. Motion. In *Handbook of Cognitive Linguistics* [Handbooks of Linguistics and Communication Science 39], E. Dabrowska & D. Divjak (eds), 527–545. Berlin: Mouton de Gruyter.

Freeman, M. & Smith, F. (eds.). *Law and Language: Current Legal Issues*, 15. Oxford: OUP.

Gibbons, J. 2003. *Forensic Linguistics*. Oxford: Blackwell.

Gibbons, J. 2011. Towards a framework for communication evidence. *International Journal of Speech, Language and the Law* 18(2): 233–260.

Hale, S.B. 2004. *The Discourse of Court Interpreting* [Benjamins Translation Library 52]. Amsterdam: John Benjamins. doi:10.1075/btl.52

Hales, L. & Gelsthorpe, L. 2012. *The Criminalisation of Migrant Women*. Cambridge: Institute of Criminology, University of Cambridge.

Hales, L. 2014. *The Language Barrier to Rehabilitation*. Cambridge: The Bell Foundation.

House of Commons Justice Committee. 2013. Interpreter and translation services and the Applied Language Solutions Contract. *Sixth Report of Session 2012–2013*. HC 245. London: House of Commons Stationary Office.

Ibarretxe-Antuñano, I. & Filipović, L. 2013. Lexicalization patterns and translation. In *Cognitive Linguistics and Translation*, A. Rojo & I. Ibarretxe-Antuñano (eds), 253–284. Berlin: Mouton de Gruyter. doi:10.1515/9783110302943

Mendoza-Denton, R. 2010. Are we born racist? Inside the science of stigma, prejudice and intergroup relations. <www.psychologytoday.com/blog/are-we-born-racist/201012/linguistic-forensics>

Ministry of Justice. April 2016. Offender Management Statistics Bulletin, England and Wales. Quarterly October to December 2015. Annual January to December 2015. <www.gov.uk/government/uploads/system/uploads/attachment_data/file/519437/offender-management-statistics-quarterly-bulletin-oct-dec-2015.pdf>

Mulcahy, L. 2011. *Legal Architecture. Justice, Due Process and the Place of Law*. Abingdon: Routledge.

National Crime Agency (NRM) <www.nationalcrimeagency.gov.uk/uk-human-trafficking-centre/national-referral-mechanism>

Quernton (Section 3). 2012. *A Gender Analysis of UK Asylum Law, Policy and Practice*. London: Asylum Aid Publications.

Roberts, F., Margutti, P. & Takano, S. 2011. Judgments concerning valence of inter-turn silence across speakers of American English, Italian and Japanese. *Discourse Processes* 48(5): 331–354. doi:10.1080/0163853X.2011.558002

Slobin, D.I. 1996. Two ways to travel: Verbs of motion in English and Spanish. In *Grammatical Constructions: Their Form and Meaning*, M. Shibatani & S.A. Thompson (eds), 195– 317. Oxford: Clarendon Press.

Slobin, D.I. 1997. Mind, code, and text. In *Essays on Language Function and Language Type: Dedicated to T. Givón*, J. Bybee, J. Haiman & S.A. Thompson (eds), 437–467. Amsterdam: John Benjamins. doi:10.1075/z.82.24slo

Slobin, D.I. 2000. Verbalized events: A dynamic approach to linguistic relativity and determinism. In *Evidence for Linguistic Relativity*, S. Niemeier & R. Dirven (eds), 107–138. Berlin: Mouton de Gruyter. doi:10.1075/cilt.198.10slo

Slobin, D.I. 2004. The many ways to search for a frog. In *Relating Events in Narrative: Typological and Contextual Perspectives* S. Strömqvist & L. Verhoeven (eds), 219–257. Hillsdale NJ: Lawrence Erlbaum Associates.

Slobin, D.I. 2006. What makes manner of motion salient? Explorations in linguistic typology, discourse, and cognition. In *Space in Languages: Linguistic Systems and Cognitive Categories* [Typological Studies in Language 66], M. Hickmann & S. Robert (eds), 59–81. Amsterdam: John Benjamins. doi:10.1075/tsl.66.05slo

Trujillo, J. 2003. The Difference in Resulting Judgments when Descriptions Use High-manner Versus Neutral-manner Verbs. Senior dissertation, University of California at Berkeley.

U.S. Federal Bureau of Prisons. 2014. Inmate Citizenship 30th August 2014. <www.bop.gov/about/statistics/statistics_inmate_citizenship.jsp>

Towards language planning for sign languages

Measuring endangerment and the treatment of British Sign Language

Jill Jones

Deaf Ex-Mainstreamers' Group Ltd / DEXperience, United Kingdom

Sign languages are well researched autochthonous, bio-culturally diverse, visio-spatial languages, both linguistically and neurologically. They confer identity and form new minorities within complex social systems. The continuity of their ethnolinguistic heritages are endangered as replacement levels fall due to the fact that sign languages are not traditionally considered to be mother tongue languages, since most deaf children are born to hearing parents who do not sign, although theoretically they are. A longstanding international discourse since 1880 when sign language was banned in education was due to the supposed effect sign languages have on spoken language acquisition for deaf children.

This ethos continues to modern day, with few parents of deaf children being informed about sign language or offered instruction, or it being used as a teaching medium for their child. This signifies the linguistic imperialism that stems from ignorance of modern research, and surdism (in which deaf people are normalised to be as hearing as possible). All the countries which have implemented sign language legislation fall short of revitalisation since there is no promotion to all parents of deaf children. This study demonstrates that the resulting extremely low number of new learners means British Sign Language (BSL) can be categorised as a severely endangered language. This paper applies spoken language planning theory and methodology to British Sign Language, taking into account the discrete political environments in the UK and history of spoken language revival. It is a rationale view of BSL, and other sign languages, as requiring immediate intervention, against the backdrop of English (and spoken language generally) being the language of oppression.

Keywords: British Sign Language, deaf, language planning, language policy, sign language/s

1.   Introduction

The topic of this contribution is the measurement of the severity of intergenerational dislocation for British Sign Language (BSL) using Fishman's seminal Graded Intergenerational Disruption Scales; his Severity of Dislocation Scales theories; expansion of, and alternatives to his theory. The treatment of disruption is addressed by the application of Fishman's Reversing Language Shift methodology to BSL. Related factors include:

–   Power discourse – normalisation or surdism, which is a longstanding international discourse about the supposedly negative effect sign languages have on spoken language acquisition for deaf children.
–   Sign language validity and lack of status.
–   Linguistic Human Rights and mother tongue definition, introducing Skutnabb-Kangas' theories.

1.1   Brief background of sign language

Langue de Signe Française was introduced in education as a French Deaf school opened in 1760, by hearing people, and more schools were established throughout Europe and the United States which developed their own sign languages, including BSL. They were sign language medium schools, until the "German method" or "oralism" began in the mid-1800s and trialled a new approach to teach deaf children to speak and to understand spoken language. The 2nd International Congress of Teachers of the Deaf-Mutes, established by the pro-oralist Pereire Society, was held in Milan in 1880. Consequently sign languages were banned in education and deaf teachers were forced to resign, resulting in linguistic genocide, which was accompanied by eugenics and a major normalisation programme.

Linguistic genocide is "prohibiting the use of the language of the group in daily intercourse or in schools, or the printing and circulation of publications in the language of the group". This was how linguistic genocide was defined in Article III(1) of the final draft of what became The Convention on the Prevention and Punishment of the Crime of Genocide (E 794, 1948) of the United Nations. The definition articles II(e) and II(b) in the present convention about what is genocide also fit indigenous and minority education. They define genocide as follows:

> Article II(e), "forcibly transferring children of the group to another group";
> and
> Article II(b), "causing serious bodily or mental harm to members of the group".

If an indigenous or minority child does not get the main part of her basic education through the medium of her own language and is forced to be in a submersion programme, with a subtractive learning environment, and if this continues for several generations, minorities are forcibly assimilated. Deaf children were physically and emotionally abused, and still are in educational establishments.

William Stokoe, an American hearing linguist, published the first paper in which the basis for the linguistic analysis of American Sign Language, and other sign languages was outlined (1960). A plethora of sociolinguist and corpus research since, led to many sign languages being legally and officially recognised, including BSL in 2003. Sign languages are accepted as biologically-natural, visiospatial, autochthonous languages. Research has found that brain processing tissue which was previously labelled *auditory tissue*, can be recruited for visual processing (Campbell et al. 200).

The evidence is that sign languages generally have managed to survive in Deaf schools due to deaf children of Deaf parents transmitting sign language to those deaf children whose parents are hearing. From approximately one hundred Deaf schools in the UK only seventeen remain, two of which are oral. The incremental advent of mainstream education began in the UK from 1951 when free hearing aids were introduced to deaf children, and when deaf children were segregated into two broad categories: (1) deaf, meaning profoundly deaf children who required assistance in learning spoken language; (2) hard of hearing children – thought to be able to hear spoken language with hearing aids. Cochlear implantation for most profoundly deaf children in the western world became common in the 1990s, and whilst hearing aids have improved, all aids have limitations (Sarant & Garrard 2014). In July 2010, the 21st International Conference on Education of the Deaf (which is the modern equivalent of the International Congress of Teachers of Deaf-Mutes), repealed the decision made in Milan 130 years earlier: "rejecting the resolutions from 1880 that banned the use of Sign Language in education for deaf children and promoted the exclusive use of speech and residual hearing called Oralism". The Conference apologised for the detriment affecting many generations of deaf people, but to date has not influenced change in deaf education.

The monolingual spoken language approach for deaf or hard of hearing children is based on the erroneous belief that early exposure to two languages would result in children growing up with poor language skills, but bilingual research has since disproved this theory. For deaf children, findings and experience from Sweden indicate, when educated bilingually, they academically perform very well, (Svartholm 2006) along with many other correlated research findings. It is obviously important that deaf children learn spoken language in all its forms as far as each individual's residual hearing permits, and develop in parallel with sign language to manifest as bilingualism. The philosophy of "informed choice" for

parents as to which language(s) their deaf child should acquire is loosely based, in the UK, on parents' legal right to choice of educational placement. There is no law in statute that maintains that parents should choose their child's language, as is assumed that every child will learn their mother tongue.

Hearing loss is the most common congenital condition, affecting 1 to 3 per 1,000 live births in the USA (Finitzo et al. 1998; Van Naarden et al. 1999). This finding is now applied in the UK. When left undetected, hearing loss of any degree, including mild bilateral (in both ears) and unilateral (in one ear) has been shown to adversely affect speech, language, and academic and psychosocial development (Schein 1996; Bess et al. 1998; Bess & Tharpe, 1986, 1988; Blair et al. 1985; Bovo et al. 1988; Brookhouser et al. 1991; Culbetson & Gilbert 1986; Davis et al. 2000, 1986; Klee & Davis-Dansky 1986; Lieu 2004; Moeller 2000; Oyler & Gross 2000; Yoshinaga-Itano et al. 1998; Richardson et al. 2010; Mason 1997; Shirin et al. 2011; Most et al. 2011). Vostanis et al. (1997) state that deaf children in mainstream education have less interaction with hearing peers, are more often rejected or neglected than their hearing peers, have a sense of loneliness, and require coping strategies. In two experiments, salivary cortisol levels and a self-rating measure were used to measure fatigue due to hard of hearing children expending more effort in listening than hearing children (Bourland et al. 1994). A high level of psychiatric disorder was also found amongst deaf children in mainstream education (61% – Hindley et al. 1994), impacting on the formation of self-image and development of psychosocial skills.

The definition of mild bilateral and unilateral hearing loss varies considerably from country to country, and even from practice to practice. The findings from DEX's literary reviews and research studies are that the majority of deaf children form a hidden group of institutionally neglected children whose needs are not being addressed appropriately. This is because the vast majority has no deaf peer group in local mainstream schools and often has minimal or no support. Not being able to hear what is being said in and outside the classroom, and the resulting fear and fatigue, can domino-effect in behaviour problems, limited concentration, feelings of inadequacy and lack of control over one's environment, impacting on self-esteem, confidence and mental health.

"Always calculate" (DEX 2004) is a major factor, especially for spoken language monolingual deaf children, encapsulating the daily struggle deaf children face in order to understand via lipreading and listening via technical aids. It is akin to doing a mental crossword without a pen or paper, having to work out what is being said in order to calculate and guess sounds and/or lip patterns that are unclear or missing. Deaf children are unable to hear all speech sounds with technical aids in group situations despite spending most of their days in groups. Classroom hearing loop systems eliminate some of the background "cocktail

party" effect as hearing aids are not fully able to help with noise discrimination, yet they only enable communication between the teacher and deaf pupil, not the deaf child's hearing peers. During the Gallaudet University's National Parent Project, 2003, 404 parents of hard of hearing children reported that their children's hearing loss was identified later than profoundly deaf children, even when family members are observant and medical care is excellent. They are not as well served as parents of deaf children, much less likely to have received counselling or information about deafness or sign language instruction, legal rights, behavioral development, school choices. Paradoxically, the positive coping skills developed by hard of hearing children sometimes contribute to their difficulties, typically communicating well in one-to-one interactions, and their good lipreading skills tend to mask the extent of their hearing loss, lulling parents and teachers into believing that they understood more than they did. Deaf children therefore, are placed in mainstream schools with no deaf peer support because uninformed parents are often denied the opportunity of their deaf child learning sign language (Meadow-Orlans et al. 2003).

The National Council for Special Education in Ireland's report (Marschark & Spencer 2009) and the report to the Finnish Government (Office of the Ombudsman for Children, Finland 2012), both outline that many profoundly deaf children utilise Communication Support Workers (CSWs), as well as different hearing devices. Children with implants are also increasingly being placed without deaf peers in their mainstream local schools. Whilst there are gaps in educational achievement compared to hearing children, DEX considers that for profoundly deaf children in a resourced mainstream school with other deaf children the ideal of social integration is somewhat more positive than for children with lesser degrees of hearing loss (DEX's Best Value Review 2004). This BVR was assessed by Boyle et al. (2007), eminent auditors of public services, as good practice, it being a deaf-led audit. Resourced classes include both deaf and hearing children who are co-taught by a general education and a teacher of deaf, and are usually supported by a CSW. Research has identified that resourced schools provide the opportunity for contact between deaf children and their hearing peers in an environment where they are not the only deaf child so they are less likely to feel lonely or isolated and do not have a lower self-esteem. Xie et al. (2014), in their literary review's conclusions reinforce co-enrollment and social skills training programs to be effective interventions for deaf children's social interaction.

1.2   Sign languages as mother tongue languages

Tove Skutnabb-Kangas (2003, 2000, 1995) redefines the concept of mother tongue, depending on four criteria:

- Origin – the language learnt first
- Identification (internal and external) – the language one identifies with or is identified by others as a native user
- Competence – the language one knows best
- Function – the language one uses most.

Skutnabb-Kangas states that sign languages are the mother tongue of deaf people, "Deaf people who grow up in hearing families can claim a sign language as their mother tongue … especially with regards to internal identification and function" (2000). She also claimed it was possible to identify with a language one does not know, or have a mother tongue without competence. Skutnabb-Kangas' theory of Linguistic Human Rights, she states, apply to sign languages (2003, 2008) just as much as to spoken languages.

Parents of deaf children in the UK (and in many other western countries) are not informed of this definition of mother tongue, and that sign languages form a complex vehicle on which deaf children's identity and culture is nurtured and shaped. This study factors in Skutnabb-Kangas' research that sign languages are the mother tongue into Fishman's Graded Intergenerational Disruption Scales, (GIDS) which is based on the traditional view of mother tongue transmission.

1.3   Applying Fishman's graded intergenerational disruption scale, (GIDS) 1991 to BSL

"The desired/discrepancy number / proportion of X-ish users within a diverse language community" is a factor to consider when measuring the health and status of a language (Mulligan 2007). It is difficult to estimate the total number of deaf BSL users since there are no reliable statistics. Turner (2009) quotes Johnston's research in Australia (2004) which found one deaf Australian Sign Language user in every 3,000 Australians. The extrapolated figure for deaf BSL users in the UK could be circa 21,000. The Office for National Statistics' census found 15,000 BSL users in England and Wales but the Deaf community considers this an underestimation. There is also an unknown figure of hearing BSL users i.e. family and professionals working with deaf people Turner (2009) terms the *Sign language community*. Whilst it is important to include all language users in measurements, it is also crucial to hold in mind that deaf people are central to sign languages, for without deaf users, sign languages have no meaning to hearing people.

Few in the community may have any sense of the impending danger outlined by Lewis (2009) and EGIDS Level 6(b). There is a common denial of rapid declines in numbers of X-ish users amongst X-ish language communities (Kaplan & Baldauf 1997; Gruffudd 2000; Fishman 2001; Nelde 2002). There appears to be a

similar evidence of denial and defensiveness of the level of severity of disruption to BSL within its language community, anxious to overstate numbers to ensure access to services.

GIDS Stage 7: Where the minority language is used by an older and not younger generation

Despite Skutnabb-Kangas' positive definition of sign languages being mother tongue languages, there is no intergenerational transmission unless parents are Deaf BSL users too. The most telling point for hearing parents of deaf children is the "perception of language usefulness" (Strubell 2001) for social advancement and the labour market (Colomer 1996). He said that young couples wanting to start a family, in the case of spoken languages, have two main factors to bear in mind: language loyalty and usefulness (also Euromosaic report in Nelde et al. 1996). For hearing parents of deaf children sign language having importance for their deaf child and for the family is paramount. It is not an understatement to say that the perceived value of BSL is low as a means to employment, and the op-portunities for many deaf entrants to the wider labour market are limited. Many bilinguals are working within the Deaf community, a choice that monolingual deaf people in English do not have.

GIDS Stage 6: Minority language is passed on from generation to generation and used in the community. Need to multiply the language in the younger generation is important to all threatened or dying languages

Lewis and Simons (EGIDS 2009) also ask Q#3: "Are all parents transmitting the language to their children?" If the answer is "*No*" this means that the "Intergener-ational transmission of L1 is being disrupted. This response would "characterize incipient or more advanced language shift".

Although difficult to judge the extent of language shift for BSL and how im-minent the threat is, there is a clear downward trend. It seems that there were 75 Deaf schools in the UK in 1982, and this had declined to 23 in 2010, reported by the Consortium of Research in Deaf Education, (CRIDE, 2011) although several more have closed since then. In the schools that remain, many have seen a decline in pupil numbers from the largest roll of 400 down to 30–50 now. Many Deaf schools now have a predominant number of deaf children with complex needs. Schools for the deaf had a major significance until the 1970s as the cornerstone of the Deaf community as BSL could be learned overtly due to the banning of BSL in education (Gregory, British Association of Teachers of Deaf website).

The number of deaf children reported by the survey conducted by CRIDE, 2014 is at least 48,125. This survey used the Freedom of Information Act 2000 so received a 98% response rate from Local Authorities and independent Deaf schools. Currently 16% have a unilateral hearing loss whilst the remainder are bilaterally deaf. 30% of deaf children have a mild loss, 31% have moderate, 11% have severe and 12% are profoundly deaf (CRIDE 2012). CRIDE, 2014, states that in England 9% of deaf children of preschool and school age, "use sign language in some form, either on its own or alongside another language". The CRIDE surveys undertaken annually since 2012 would extrapolate a figure of 3,500 deaf children using BSL in the UK.

Approximately 95% of deaf children are born to hearing parents who do not usually know sign languages (Mitchell & Karkchmer 2002) compounded by the stress of having a deaf child (Calderon & Greenburg 1997, 2000; Moeller 2000). Only 5% of deaf children are able to learn BSL naturally from their Deaf parents, assuming that all their parents sign. The teaching of BSL to parents even where provided is only usually of 6 weeks' duration and parents wanting to learn more have to pay to attend local adult education classes.

Fishman's GIDS Stage 6, therefore, demonstrates the distressing lack of the integral stage for BSL transmission because there has never been opportunity for BSL to be a natural mother tongue, except for the tiny minority of Deaf parents who have deaf children. This is where the term "inter-professional transmission" instead of intergenerational transmission and Governments' adoption of a "language in loco parentis" role could be brought into play as a means of describing how to apply Fishman's GIDS Stage 6 to sign languages, vital as this stage is to all language maintenance.

Language planners must address how to enable parents to quickly accept their deaf child's need for a mother tongue alongside English as being integral to BSL's maintenance. As the Deaf group of native users decreases over decades, the demand for other BSL professionals is likely to dwindle and die out, i.e. interpreters.

GIDS Stage 5: Schools for literacy acquisition, for the old and the young, and not in lieu of compulsory education

This level suggests the need to support literacy movements in the minority language, particularly where there is no government support. On-line or video signing is the "written" version of spoken languages. Within the education system BSL currently has no formal system to reproduce itself in the same way as literacy acquisition in spoken languages.

GIDS Stage 4(a): Schools in lieu of compulsory education and substantially under Xish curricular and staffing control

Almost all schools attended by deaf children are managed by hearing people most of whom do not sign either within the state system or in independently run Deaf schools, (the requirement by the Department for Education for teachers of deaf is Level 1 in BSL, equivalent to conversational level). The amount of time studying the language and the low quality of communication support, and the few deaf peers they have in school, is often resulting in deaf children leaving school with a much lower level of fluency in BSL than hearing adults who have learnt to sign to interpreting level (Turner 2009). Without binding educational linguistic human rights, especially a right to mainly mother tongue-medium education in state schools, with good teaching of a dominant language as a second language, given by competent bilingual teachers, most indigenous peoples and minorities have to accept subtractive education through the medium of a dominant/majority language. Subtractive teaching subtracts from the child's linguistic repertoire (Nicholaisen et al. 2005).

GIDS Stage 4(b): State or public schools for Xish children, providing some instruction via Xish, but substantially under Yish curricular and staffing control

Whilst there are fifteen BSL-using Deaf schools, all are not BSL-medium since many use a combination of English and BSL simultaneously (a sign system) and this "total communication" is widely used through resourced mainstream schools. Some Deaf schools are co-located with mainstream schools, particularly at secondary level. For all deaf children exposed to BSL, English is taught as a subject and used for all written instruction and examinations in other subjects (except in Scotland where pupils can choose to take their exams in BSL, receiving questions in BSL and signing their responses to video camera). Deaf Instructors or BSL Tutors in schools are Deaf adult BSL users, many of whom are not qualified to teach children, so the role is low paid and often of low status. BSL Tutors can qualify for adult teaching via vocational training routes.

GIDS Stage 3: The local/regional (i.e. non-neighbourhood) work sphere, both among Xmen and Ymen

This stage relates to the use of BSL in the work sphere outside the Deaf community, involving interaction between Deaf individuals and hearing spoken language communities. Unless BSL users work with other BSL users there is no BSL usage amongst the working population due to there being widespread ignorance about this indigenous language. A high level of interest was raised amongst hearing

people in learning BSL as a leisure activity at adult learning centres when the UK government subsidised course fees, which in some cases encouraged learners to become BSL interpreters and interest decreased when centres started charging in full. BSL is used in businesses established by Deaf entrepreneurs, such as BSL teaching and Deaf Equality training companies, or where they manage interpreting agencies, and in education with signing deaf children. In the workplace where BSL is not used, Deaf people are entitled to government Access to Work funding which purchases the services of BSL/English interpreters or CSWs. This funding can be dependent on employers' ability to financially contribute, and because the scheme is poorly promoted, many Deaf people do not receive this support. This is the stage of greatest concern to deaf adult BSL users, in order to access non-BSL work environments and services. The establishment of a fluent workforce must be an integral part of a BSL language plan.

2.   Local/regional mass media and governmental spheres

Public and private services are goods and services targeted by the UK's Equality Act 2010 in order to enable access for deaf and disabled people. The Act, however, is not making sufficient impact, for whilst there is best practice, many public and private organisations do not comply, largely due to BSL being undervalued and unpromoted. The thrust of this legislation for disabled people (2) (20) is the "duty to make reasonable adjustments", which can prevent access if the provider deems it unreasonable to offer this. The Deaf community wants sign languages to be taken out of the disability arena and for them to be treated in the same way as minority spoken languages, but it cannot make this leap as long as the overriding view is of them being disabled. So, whilst calling for legislation to strengthen services in and via BSL, the Deaf community's call for BSL's legal recognition naturally focuses on the increase and quality of interpreters and wider range of access (i.e. Strubell's Catherine Wheel 1999 – "more demand for goods and services in the language" leading to "more supply and consumption").

2.1   Education, work sphere, mass media and governmental operations at higher and nationwide levels

This is thought to be the natural level on which to focus language planning attention, but this agenda for the Deaf community, as with other minority language revitalisers, is virtually impossible. It is true that, within the Deaf community, as Fishman (2001:5) posits: "RLS is an activity of minorities, frequently powerless,

unpopular with outsiders and querulous amongst themselves". The UK legal system has, over time, recognised its duty to provide BSL interpreters. Lawyers often in court do not have the financial resources to pay for interpreters, deaf people are not permitted on jury service and to our knowledge there have been no legal challenges to date from deaf people to ensure their linguistic and human rights.

In further and higher education in the UK CSWs are usually not fully qualified and registered interpreters. There is government funding available to deaf students for this service but is often insufficient to cover costs. In mass media translation to BSL is hardly conceived of, with isolated pockets of Deaf-led companies translating public service information, but few filmed parliamentary debates are translated online. There are no deaf BSL using Ministers of Parliament, several deaf people have become local councillors but with no funding for interpretation at constituency surgeries, nor any funding for access to local party political activities.

2.2   Applying other language planning research to BSL

Other language planning theories and guidelines incorporated in this study are the Language Vitality and Endangerment Methodological Guidelines (LVE) (UNESCO 2003); Strubell's Catherine Wheel, 1999; Lewis and Simon's Expanded GIDS (2010), and spoken language "success stories". UNESCO's LVE Framework (2009) states as one of the degrees of endangerment:

> Severely endangered – the language is spoken by grandparents **and older generations** [my emphasis]; while the parent generation may understand it, they do not speak it to children or among themselves.

If circa 3,000 deaf children are using BSL in the UK, with likelihood of a decreasing population of deaf children and fewer being exposed to BSL in the near future, it can be stated from Fishman's and LVE's criteria that BSL is severely endangered.

3.   Reversing language shift and BSL

3.1   Consideration of reversing language shift and its relevance
to sign languages

Having measured BSL against Fishman's Severity of Dislocation Scale and LVS, this study has established the importance of using these measurements and that, according to these criteria, there is severe dislocation and endangerment. This

section focuses on how to redress the severity of its shift using Fishman's Stages of Reversing Language Shift: Severity of Intergenerational Dislocation (2001: 466; Figure 19.1).

3.2   RLS to attain diglossia

Continuity of research should be funded to support the divergent varieties which are shaping BSL into a rich language that is fit for purpose to incorporate educational, medical and legal corpus. Research outcomes must be centrally controlled to advance BSL, as with Welsh language research projects, and also focus on BSL acquisition.

Stage 8. Reconstructing Xish and adult acquisition of XSL (and)
Stage 7. Cultural interaction in Xish primarily involving the community-based older generation

A lexical database documenting 50,000 signs from four regions from the BSL Corpus data has been transformed into an online dictionary, BSL SignBank, as part of work on language documentation and language change by DCAL. This is a reference grammar of BSL, discussing key characteristics of the phonology and morphosyntax of BSL, and is an ongoing project. The community-based older generation of Deaf people constitutes much of the BSL teaching force, either in the community or in schools, with a grave risk of this group not being replaced by younger Deaf people to maintain the current level of BSL teaching. The seven current GCSE awarding bodies in England, Wales and Northern Ireland and the Scottish Qualifications Authority should offer deaf and hearing children and adults BSL examinations, i.e. national curriculum required by the New Zealand Sign Language Act.

Stage 6. The intergenerational and demographically concentrated home-family-neighbourhood-community

To address this risk-laden stage the numbers of deaf children learning fluent BSL to enable them to teach it when older must be kept to an optimum level which is, ideally, the inclusion of all deaf children with mild, moderate and severe losses, an additional 72% (CRIDE 2012). The current practice of offering BSL to parents of profoundly deaf children will mean numbers will either stay at this level or continue to decline over coming decades. For deaf children BSL should be a mandatory, accredited subject to examination level in statutory education and further

and higher education. As a minimum standard they should study BSL as a subject, with an option for hearing children, creating a potential pool of professionals.

An organisation commissioned by the Welsh Language Board, Twf, promotes Welsh as an essential language to all pregnant mothers and parents of newborn babies in Wales. This grassroots promotion of Welsh has been highly successful, and is a model that DEX considers could easily fit BSL. This approach encourages parents to opt for Welsh–medium education, as the most effective means. The numbers of Welsh children attending Welsh-medium schools is slowly increasing due to this organisation's work, which includes giving prospective and new parents quality information, teaching and support. Twf has a network of field officers located throughout Wales, conveying to parents and the general public the cultural and economic advantages of raising children bilingually. An evaluation of Twf:

> has grown rapidly in a short period of time and has been extremely successful in transforming the abstract concept of family bilingualism into a concrete message with which the target audience can identify.
>
> (Edwards & Pritchard Newcombe 2005: 135)

This programme could easily be replicated at the point of identification of deafness at the national Newborn Hearing Screening Programme or at later stages. On this care pathway parents of deaf children are advised by medical and teaching professionals (usually hearing people). As in Scandinavia, free BSL classes for parents and extended family members, for up to a year, with state reimbursement to employers for employees' time lost from work should be a legal requirement.

Stage 5. Schools for literacy acquisition, for the old and for the young, and not in lieu of compulsory education

The development of BSL on-line teaching and translation services requires systemic overview, registration and the sharing of materials and resources in a governmental BSL Board Library for BSL teachers, learners and the Deaf community. More cost effective than the printed word, resources can be produced at relatively low cost compared to paper documents.

Stage 4(b). Public schools for Xish children, offering some instruction via Xish, but substantially under Yish curricular and staffing control

Sweden established Swedish Sign Language (SSL) as a first language for profoundly deaf children in its Education Act 1989, and established a teacher trainer course for deaf SSL users to become qualified teachers of SSL. This has given them parity to other teachers in the state system. An important component of a BSL

language plan is that this concept is legally adopted in the UK for all future BSL initial teacher training. Teachers of deaf children working directly with BSL users should have level 3 or 6 qualification in BSL and Education Interpreters (currently CSWs) should be fully qualified interpreters with specialism in education.

Stage 3. The local/regional (i.e. non-neighbourhood) work sphere, both among Xmen and Ymen

A BSL language plan initiative must include all RLS stages since conferring the same rights as hearing citizens of the UK, and placing BSL in the power function of governmental control, will support an improved quality of life for Deaf BSL users. There should be opportunities for hearing people to learn BSL so that there is direct, albeit more low level, communication in local/regional work sphere, including front line staff of public and retail services. BSL Standards mandated for by legislation should target public and private sectors on how to efficiently fund and acquire registered BSL/English interpreters. Their training, work conditions, salary, ratio per Deaf users and quality assurance would be consistently applied by a BSL Board according to statute.

Stage 2. Local/regional mass media and governmental spheres

An appropriately worded BSL Act would require more BSL translation of television and on-line programmes and information, throughout peak viewing times, and more special programmes for Deaf people. The Deaf Broadcasting Council, now the BSLBT, has done sterling work and should advise the BSL Board.

Stage 1. Education, work sphere, mass media and governmental operations at higher and nationwide levels

According to this criteria there should be discrete legislation that primarily mandates for the increase in the number of users of BSL, as well as raising its status as an equal language to English, Welsh and Gaelic in the UK. It is proposed that the management of this is by the governmental cross-departmental establishment of a statutory BSL Board of language planning experts, with ongoing involvement with the Deaf community to give the Board powers to monitor and evaluate, using BSL Standards.

3.3   Welsh language planning

As part of its BVR, 2004, DEX researched the Welsh language and updated its findings for this contribution. The Welsh language is perhaps the most celebrated example of successful language planning, largely through the sustained campaign by the Welsh Language Society, some of whose members were internationally acclaimed linguists. Over the course of the twentieth century the monolingual Welsh population all but disappeared, but a small percentage remained at the time of the 1981 census. The Welsh language has been statutorily promoted through the national curriculum in Wales as a mandatory subject to be learnt within Wales, enacted through the Education Reform Act 1989. The Welsh Language Act, 1993 ensured the effective promotion and monitoring of Welsh via the Welsh Assembly overseeing the compilation of Welsh Language Schemes by public authorities to demonstrate actions taken in the delivery of Welsh in their services. The Welsh Language (Wales) Measure 2011 includes provisions about the official status of the Welsh language, and states that the Welsh language should not be treated less favourably than the English language in Wales. The Measure established the Welsh Language Commissioner, which replaced the Welsh Language Board, who has powers to decide if a public organisation has failed to carry out its statutory duties and makes recommendations how the organisation should take to rectify the situation, and investigate complaints from Welsh speakers. The Measure continues to develop Standards, formerly called Welsh Language Schemes and created a Welsh Language Tribunal (Welsh Ministers' report 2014). The Welsh Assembly's Welsh Language Strategy (2012–2017) states "the results of the 2001 Census showed that 20.8 per cent of the population of Wales was able to speak Welsh (582,400 people) compared to the 1991 Census (18.7 per cent and 508,100 people). The Welsh Assembly's research suggests focussing Welsh language marketing and promotional approaches on younger age groups, maximising the potential of technology such as the internet and new social media, and improving the accessibility to, and relevance of, available Welsh-language media and resources.

3.4   Gaelic language planning

Gaelic is not a mandatory language in Scotland's education system. Despite Gaelic language RLS efforts (Dorian 1997b, 1980, 1981) the census of 2011 in Scotland indicated that a total of 57,375 people (1.1% of the Scottish population aged over three years old) in Scotland could speak Gaelic at that time, with a decline of 1,275 Gaelic speakers from 2001. A total of 87,056 people in 2011 had some ability in Gaelic compared to 93,282 people in 2001, a reduction of 6,226 users. The

Gaelic Language (Scotland) Act 2005 required a Gaelic Language Board (Bòrd na Gàidhlig) and Gaelic Language Plans and the Bòrd to regularly produce a National Plan. The Scottish Parliament's National Plan for Gaelic (2014) states the Bòrd na Gàidhlig identified its priority action areas, including language, status and corpus, promote Gaelic-medium education, working with local authorities, media and in the arts, tourism, heritage and recreation sectors, quality of Gaelic translations, and research. Gaelic's policy history includes legislation with Gaelic provisions: the Education (Scotland) Act 1980, the Standards in Scotland's School etc. Act 2001, the Schools (Consultation) (Scotland) Act 2010 and the Gaelic Language (Scotland) Act 2005. This model is of greatest interest to the Deaf community in the UK since the BSL (Scotland) Act 2015 was enacted based on spoken Gaelic legislation, and proposed that there is a Minister responsible, a National Advisory Group is established and BSL Plans are developed by the Scottish Parliament and public authorities in Scotland, with extensive consultation and the greatest ever support from the UK Deaf community. DEX has observed that the the Scottish Act merely perceives BSL as an access tool, reinforced by De Meulder, 2014. DEX has outlined to the Scottish Parliament and Westminster about the damaging limitations of not factoring in a legal substitute for the lack of natural intergenerational transmission of BSL into any proposed legislation in the UK.

4.   Policy and conventions

How can linguistic theory be applied so there is a formalised "linkage system of adult functions and institutions that are *prior to and preparatory for* schooling for children" (Fishman 2001: 15). In the case of sign languages there has to be a clear distinction drawn between the use of (a) language planning, (b) language policy and (c) generalised legal policy using legislation aimed at disability or amending existing legislation/ regulations for sign languages. Sign language legislation generally falls into the latter category as is evidenced by international sign languages' legislation or regulations. GIDS and UNESCO LVE Framework, 2009 should be firmly embedded into initiatives to safeguard sign languages.

4.1   Legislation and UN conventions

The following investigation as to international legislation of sign languages and UN Conventions indicates that not all RLS stages have been brought to play though there are steps to progress.

4.1.1   *UNESCO Salamanca Statement and Framework for Action on Special Needs Education (1994)*

In 1994 more than 300 participants representing 92 governments and 25 international organisations met in Salamanca, Spain, to further the objective of education for all by considering the fundamental policy shifts required to promote inclusive education, particularly for those with special educational needs. The Statement adopts the principle of inclusion by working towards "schools for all" – institutions including everyone, celebrating differences and responding to individual needs.

II (A) (21) states:

> The importance of sign language as the medium of communication among the deaf, for example, should be recognized and provision made to ensure that all deaf persons have access to education in their national sign language. Owing to the particular communication needs of deaf and deaf/blind persons, their education may be more suitably provided in special schools or special classes and units in mainstream schools.

4.1.2   *The UN Convention on the Rights of Persons with Disabilities, 2008*

Many of the world's governments have ratified this convention, and committed to following its recommendations which include each member states' recognition and promotion of their country's sign language. The UK's Protocol ratification date was in 2009, and a government report has been completed stating what UNCRPD targets have been achieved and gaps identified. DEX was asked by the Equality and Human Rights Commission and the Office of the Deputy Prime Minister to submit a Shadow Report to the government's response to be considered shortly. DEX's Shadow Report gives a broad outline of BSL's decline and it is submitting the findings of this study to the UNCRPD Committee in our Alternative Report, updating the Shadow report.

4.1.3   *European Parliament*

In 1988 the European Parliament Resolution on Sign Languages for Deaf People was called on by the European Parliament, requiring the Commission to make a proposal to the Council concerning the official recognition of the sign language in each member state, calling upon member states to abolish any remaining obstacles to the use of sign languages. The wide ranging proposals covered teaching, training of teachers and interpreters, compilation of dictionaries, broadcasting, and technology. The World Federation of Deaf People has worked in coordination with the EUD on this matter.

4.1.4   *European Union of the Deaf*

Some Deaf members of the European Union of the Deaf (EUD) are also members of their own country's European Union councils, and this has assisted in raising the profiles of sign languages in Europe, ie. the EUD report of 2005 was the first account of the state of play for sign language status in individual countries.

4.1.5   *European Charter for regional or minority languages*

Deaf communities in Europe appealed to the Council of Europe for sign languages' inclusion in its European Charter for Regional or Minority Languages, 1992, to persuade it to explicitly mention sign languages as part of its remit. The Parliamentary Assembly of the Council of Europe issued its Recommendation 1598 in 2003, that "the Committee of Ministers devise a specific legal instrument on the rights of sign language users". Subsequently the Parliamentary Assembly of the Council of Europe issued Motion for a Resolution, Doc. 10636, 2005, (6) "The Assembly therefore calls on the Committee of Ministers to proceed without delay to draft and consult on such a legal instrument that specifically delivers the following:

i.   Legal recognition of the relevant sign languages within each member state
ii.   The right of parents of deaf children to be advised as to the role that sign language could play in their lives
iii. The right of the deaf child to be taught with the use of sign language if they so wish.

Apart from the fact that this Motion was never followed through, in it there are two major flaws, as it has not applied basic language planning theory: that parents are the missing key to transmission, and the phrase "could play in their lives" has not emphasised the importance of promotion. Item (iii) did not consider the age of the deaf child expressing a wish to learn sign language, and that leaving it until the child is capable of making a judgement delays the opportunity for the language to be acquired at an early age.

Internationally, resources available to minority languages, too, are "often quite meagre … with no outside support of any operational significance to fall back upon" compared to those for the mainstream host languages (Fishman 2001:13). The current global economy is such that deemed luxuries such as services for minorities are being severely cut, not just language minorities. Ladd (2003) states that when there are cutbacks in service levels due to economic constraints, low priority is always given to deaf people's needs.

4.1.6   *Legal status of minority languages as part of RLS*

The legal recognition of sign languages is one of the major concerns of the international Deaf community. Each country has its own constitution or parliamentary system which makes it difficult to find a standard route to ensure the same legal status for all sign languages. In some countries, the national sign language is an official state language, whereas in others it has a protected status in certain areas such as education. However, symbolic recognition or legislation is no guarantee even that sign language users' quality of lives will be bettered. This section considers how legal status has affected the development of minority languages in various parts of the world, and whether state overall control has conferred power, language spread, corpus and acquisition planning. With respect to sign languages different policies adopted by countries throughout the world have:

i.   recognised their sign language is a *bona fide* language of the country as an official language by amending the constitution or by other means
ii.   recognised the sign language as a language per se
iii.   mentioned sign language in education law or policy
iv.   written their sign languages into other existing laws
v.   established a new law to enshrine the country's sign language(s).

Each country's Deaf community and its allies have actively pursued many routes applicable to each passage of legislation or official declarations. In some cases countries have more than one sign language which have been mentioned in law or regulations:

Belgium – French Community, 2003 (French-Belgian Sign Language)
–   Flanders (Flemish Sign Language) 2006
–   Wallonia (French-Belgian Sign Language) 2003
Spain – Spanish Sign Language and Catalan Sign Language, 2007, 1994, 2006.

The following countries have used education acts to legislate for sign language acquisition; specifically legislated for its sign language; mentioned their sign language in other legislation; and recognised it; (though perhaps not giving the country's sign language its name):

Austria 2005; Brazil 2005; Bulgaria 2012; Canadian Providences: Manitoba 1988, Alberta 1990, Ontario 2007 (American Sign Language); Cyprus 2006; Czech Republic 1998, 2008; Denmark 2005; Estonia 2007; Finland 1995, 2003; France; Germany 2002; Greece 2003; Hungary 2011; Iceland 2004, 2011; Ireland 1998; Latvia; Lithuania; Macedonia; Netherlands 2003; New Zealand 2006; Norway, 1999; Poland 2011; Portugal; Romania; Slovak Republic 1995;

Slovenia; Sweden 1981; Switzerland; Thailand 1999; South Africa 1994/1996; Uganda 1995; UK 2003; Uraguay 2001; Venezuela 1999.

(Wheatley & Pabsch, European Union of Deaf 2012)

4.2.1   *UK recognition of BSL*

The Department for Works and Pensions (DWP) of the UK government announced BSL recognition on 18 March 2003 (Hansard):

> "The Government recognise that British Sign Language (BSL) is a language in its own right regularly used by a significant number of people …. The Government understand that people who use BSL want their language to be protected and promoted in the same way some minority languages are by the Council of Europe's Charter for Regional or Minority Languages. The Council is considering how that might be achieved for indigenous sign languages. The Government will give careful consideration to any proposals which the Council might make. The Government have already taken action to improve access to BSL, for example by identifying situations where it might be reasonable for employers and service providers to engage the services of a BSL/English interpreter. The Government will be funding a discrete programme of initiatives to support this statement".

£1.5 m funding was allocated for this programme to be used to increase awareness of "the communication needs of deaf people who use BSL and increasing opportunities for people to study BSL at a professional level." The funding allocated for recognition was used for training of BSL/English interpreters, aspiring to and modelled on Finland's greater ratio of interpreters per Deaf people. The recognition of BSL is not in statute. The UK has three laws mentioning sign language and interpreters: Police and Criminal Evidence Act 1984; Mental Capacity Act 2005; Equality Act 2010. Following the 10th year anniversary of the DWP announcement, there has been a re-emergence of interest within the Deaf community calling for a BSL Act, with over 11,000 deaf people clicking on to a Facebook page and lobbying from Deaf organisations for action.

4.2.2   *Commentary on legal status and recognition of sign languages internationally*

Most of the countries mentioned above have treated their sign language(s) as access tools to increase the number of interpreters to access mainstream services or to promote them to service providers. The New Zealand Sign Language (NZSL) Act, 2006, declared NZSL a voluntary subject in the national curriculum, and an official language of New Zealand, giving it equal status to Maori and English. A review by the New Zealand Office for Disability Issues, 2011, states: "neither the Māori Language Act 1987 nor the NZSL Act specifies what designation as

an 'official language' means". A considerable number of submitters to the Review observed that Deaf children in mainstream education have poor access to NZSL. The New Zealand Human Rights Commission reported on the state of NZSL since the Act and recommends in its report:

> The Ministry of Education plays a key role in enabling not only access to NZSL but also its maintenance and promotion. Therefore, the Commission's strongly preferred option is that the Ministry is the agency charged with leading the development of a NZSL Statutory Board.

The use of the term "deaf" in all countries' legislation is not clear unless stated, i.e. Norway 1999 legislates for profoundly deaf children. DEX visited Sweden and Norway as part of its BVR and there it was ascertained that hard of hearing children were not encouraged to be part of the bilingual programme, although many of the hard of hearing young people DEX interviewed wished that they had been included. There was a special school for hard of hearing children at the time of the BVR visit in 2002 which was not a Swedish Sign Language-medium school. Denmark is the only country that actually qualifies in law that the teacher determines whether or not the deaf child can use Danish Sign Language although this is standard practice throughout the world. Legislation is seen as a major victory for longstanding campaigners, but until effective RLS efforts are applied, this has to be a hollow victory.

Legislation in Greece states Greek Sign Language is the language of deaf and hard of hearing pupils and teachers need to know sign language. Thai Sign Language is used by 20% of deaf students in education, a significant figure. Macedonian Sign Language's official recognition does not state that users should be deaf, "a natural way of communication between the people" which may promote language spread amongst the range of deaf, as well as hearing, children.

Branson and Miller (2008) argue that national sign languages should not be the only sign languages to be addressed by policy-makers, but that localised sign languages should also be included: i.e. Canada has Langue de Signes Quebecoise, Inuit Sign Language and American Sign Language within its multilingual remit. The Austrian Federal Government has proposed its National Action Plan for the implementation of the UNCRPD, dedicating a full chapter to sign language. Whilst the Austrian Deaf Association welcomes this, it is concerned that this chapter only addresses the public administration and funding area, and "*that securing communication via sign language interpretation alone is not the way forward.*" In Ireland a bill for the legal recognition (Recognition of Irish Sign Language for the Deaf Community Bill 2013) was defeated in January 2014, and rejected by 3 votes (24–21) at its second reading. Proposed by the opposition party,

the Minister of State for Health opined that "this Bill is putting the cart before the horse", as it was considered that the bill had not been properly thought through.

Despite Finland's efforts to successfully increase the numbers of interpreters, with just 5,300 users, the endangered status of Finnish Sign Language has been acknowledged in 2011 by the Language Policy Programme and a Memorandum by the Ministry of Justice. The Human Rights Council Resolution 22/3, with Answers by the Finnish Association of the Deaf in 2013, highlighted the fact that, despite recognition in legislation, Finnish Sign Language is threatened. The Finnish government passed The Sign Language Act in March 2015, referring to Finnish Sign Language and Finland-Swedish Sign Language. A signer is defined in the Act as a person "whose own language is sign language" – a definition to include both deaf and hearing people.

5.   Concluding recommendations

The summary of overall findings of this study are that, when compiling BSL language policy in the form of legislation it is, therefore, crucial to incorporate the language planning field of enquiry, which also includes other linguistic theories and practice, all informing each other. Because language policy has traditionally addressed spoken language, simply focusing on a generic or social policy approach is detrimental to the future of BSL and to its users, and viewing BSL simply as an access tool to mainstream services, although, granted it is one, threatens its existence even more.

Fishman (2001) states that language planning has to take into account the high power stages above GIDS stage 6, which is the establishment. Threatened languages such as BSL need to be given a secure functional niche. However, although giving BSL status will give it security, other RLS efforts must be applied consistently.

The practice of focusing on hearing levels of deaf children as a gateway to learn BSL must end. Not only is this detrimental to the health of the language and to deaf children's wellbeing but it is a fundamental breach of linguistic and human rights. If the state acts as intergenerational transmitter, and undertakes its duty of safeguarding BSL in a *linguistic loco parentis function* for deaf children and the wider Deaf community, this will demonstrate a will to effectively BSL language plan.

DEX has determined that the Welsh Language model is applicable spoken language law to act as the structure for effective BSL legislation. DEX has the support of some of the linguists who attended the LAUD Symposium (2014).

Colin Baker, internationally renowned expert on bilingualism, and Meirion Prys-Jones, Chief Executive of the Network to Promote Language Diversity, who was also the Chief Executive of the former Welsh Language Board, have advised and supported DEX's work. DEX's suggested RLS approach is based on its original BVR recommendations, strengthened by this language planning study and, thus proposes RSL treatment: a BSL Board consisting of committed experts in the field of language planning, recommended to the government for appointment, in line with the Welsh Language Board's development. The BSL Board should be a government appointed advisory body, with the powers to make recommendations to the relevant Secretaries of State and with a governmental method of appeal against decisions. In this hierarchical sense it should follow the Welsh Language revival model.

Parents of deaf children have a right to know what the best outcomes are for their deaf child, based on a huge body of research. This study has brought about a realisation that the individualistic wellbeing and safeguarding approach to deaf children should become a collective linguistic rights issue.

Two quotes emphasise why there is need for a well perceived language plan for BSL, and as a model for other sign languages: "When my teacher signs I hear everything she says" (The Swedish National Association for Deaf, Hearing-Impaired and Language-Impaired Children's website).

Bernard Spolsky, Professor Emeritus at the University of Tel Aviv, has kindly read through DEX's research. He gave his verdict: "BSL is severely endangered, and deaf children are also severely endangered."

References

Austrian Sign Language recognition. <http://www.oeglb.at#>

Bess, F. & Tharpe, A. 1986. Case history data on unilaterally hearing-impaired children. *Ear and Hearing* 7: 14–19. doi:10.1097/00003446-198602000-00004

Bess, F. & Tharpe, A. 1988. Performance and management of children with unilateral sensori-neural hearing loss. *Scandinavian Audiology. Supplementum* 30: 75–79.

Bess, F.H., Dodd-Murphy, J. & Parker, R.A. 1998. Children with minimal sensorineural hearing loss: Prevalence, educational performance, and functional status. *Ear and Hearing Journal* 19: 339–354. doi:10.1097/00003446-199810000-00001

Blair, J., Peterson, M. & Viehwed, S. 1985. The effects of mild sensorineural hearing loss on academic performance of young school-age children. *The Volta Review* 87: 87–93.

Bourland, C.H. & Tharpe, A.M. 1994. Listening effort and fatigue in school-age children with and without hearing loss. *Journal of Speech and Hearing Research* 37: 216–226. doi:10.1044/jshr.3701.216

Bovo, R., Martini, A., Agnoletto, M., Beghi, D., Carmignoto, D., Milani, M. & Zangaglia, A.M. 1988. Auditory and academic performance of children with unilateral hearing loss. *Scandinavian Audiology. Supplementum* 30: 71–74.

Branson, J. & Miller, D. 2008. National Sign Languages and Language Policies. In *Encyclopedia of Language and Education, 2nd edition, Volume 1: Policy and Political Issues in Education*, Stephen May & Nancy H. Hornberger (eds), 151–165. Berlin: Springer.

Brookhouser, P.E., Worthington, D.W. & Kelly, W.J. 1991. Unilateral hearing loss in children. *Laryngoscope* 101: 1264–1272. doi:10.1002/lary.5541011202

BSL Corpus Project. <www.bslcorpusproject.org/data/>

Calderon, R. & Greenberg, M. 2000. Challenges to parents and professionals in promoting socioemotional development in deaf children. In *The Deaf Child in the Family and at School*, P.E. Spencer, C.J. Erting & M. Marschark (eds). New Jersey: Lawrence Erlbaum Associates.

Campbell, R., Brammer, M., Calvert, G., David, A., MacSweeney, M. & Woll, B. 2000. Language, deafness and the brain: fMRI findings. *International Journal of Psychology* 35(3): 225–226.

Colomer, J.M. 1996. *La utilitat del bilinguisme*. Barcelona: Edicions 62.

Consortium of Research in Deaf Education. 2012. *Survey*. <www.batod.org.uk/index.php?id=/resources/survey/CRIDE2012>

Culbertson, J.L. & Gilbert, L. E. 1986. Children with unilateral sensorineural hearing loss: Cognitive, academic, and social development. *Ear and Hearing* 7: 38–42. doi:10.1097/00003446-198602000-00007

Davis, J.M., Elfenbein, J., Schum, R. & Bentler, R.A. 1986. Effects of mild and moderate hearing impairments on language, educational, and psychosocial behavior of children. *Journal of Speech and Hearing Disorders* 51: 53–62. doi:10.1044/jshd.5101.53

Davis, A., Reeve, K. & Hind, S.B.J. 2000. Children with mild and unilateral hearing loss. In *A Sound Foundation through Early Amplification 2001: Proceedings of the Second International Conference*, R.C. Seewald & J.S. Gravel (eds), 179–186. Bury St. Edmunds: St. Edmundsbury Press.

Deaf Ex-Mainstreamers Group. DEXperience. 2004. *Deaf Toolkit: Best Value Review of Deaf Children in Education, from Users' Perspective*. Delta Press.

Deafness Cognition and Language Research Centre University College London. <http://www.ucl.ac.uk/dcal>

De Meulder, M. 2014. Whose Recognition? – The British Sign Language (Scotland) Bill. Ms. <https://www.researchgate.net/publication/269094240_Preprint_Whose_%27recognition%27_the_British_Sign_Language_Scotland_Bill>

Dorian, N.C. 1980. Language shift in community and individual: The phenomenon of the laggard semi-speaker. *International Journal of the Sociology of Language* 25: 85–94.

Dorian, N.C. 1981. *Language Death: The Life Cycle of a Scottish Gaelic Dialect*. Philadelphia PA: University of Pennsylvania Press.

Edwards, V. & Pritchard Newcombe, L. 2005. Language transmission in the family in Wales: An example of innovative language planning. *Language Problems & Language Planning* 29(2): 135–150. doi:10.1075/lplp.29.2.03edw

European Union. Resolution on sign languages for the deaf. Official Journal C 187. 18 July1988, 0236.

European Union of Deaf. 2005. The status of sign languages report.

Finitzo, T., Albright, K. & O'Neal, J. 1998. The newborn with hearing loss: Detection in the nursery. *Paediatrics* 102: 1452–1459.

Fishman, J.A. 1991. *Reversing Language Shift: Theory and Practice of Assistance to Threatened Languages*. Clevedon: Multilingual Matters.

Fishman, J.A. 2001. *Can Threatened Languages be Saved?* Clevedon: Multilingual Matters.

Gruffud, H. 2000. Planning for the use of Welsh by young people. In *Language Revitalization: Policy and Planning in Wales*, C.H. Williams (ed.), 173–207. Cardiff: University of Wales Press.

Hansard. UK Government. UK Parliamentary records. 2003. 18 March 2003. <www.publications.parliament.uk/pa/cm200203/cmhansrd/vo030318/wmstext/30318m02.htm>.

Hindley, P.A., Hill, P.D., McGuigan, S. & Kitson, N. 1994. Psychiatric disorder in deaf and hearing impaired children and young people: A prevalence study. *Journal of Child Psychology and Psychiatry* 35(5): 917–934. doi:10.1111/j.1469-7610.1994.tb02302.x

Johnston, T. 2004. W(h)ither the deaf community? Population, genetics, and the future of Australian sign language. *American Annals of Deaf* 148(5): 358–375. doi:10.1353/aad.2004.0004

Kaplan, R.B. & Baldauf, R.B. 1997. *Language Planning from Practice to Theory*. Clevedon: Multilingual Matters.

Klee, T.M. & Davis-Dansky, E. 1986. A comparison of unilaterally hearing-impaired children and normal-hearing children on a battery of standardized language tests. *Ear and Hearing* 7: 27–37. doi:10.1097/00003446-198602000-00006

Ladd, P. 2003. *Understanding Deaf Culture: In search of Deafhood*. Clevedon: Multilingual Matters.

Lewis, M.P. & Simons, G.F. 2010. Assessing endangerment: Expanding Fishman's GIDS. *Revue Roumaine de Linguistique* 2: 103–120.

Lieu, J.E. 2004. Speech-language and educational consequences of unilateral hearing loss in children. *Archives of Otolaryngology and Head and Neck Surgery* 130: 524–530. doi:10.1001/archotol.130.5.524

Marschark, M. & Spencer, P.E. 2009. *Evidence of Best Practice Models and Outcomes in the Education of Deaf and Hard-of-Hearing Children: An International Review*. A report commissioned by the National Council for Special Education (Ireland).

Mason, D.G. 1997. Mainstream education and deaf students. *Canadian Annals of Education for Deaf and Hard of Hearing* 23.

Meadow-Orlans, K.P., Mertens, D.M. & Sass-Lehrer, M.A. 2003. *Hard of Hearing Children: Still Overlooked*. Odyssey: Winter.

Meadow-Orlans, K.P., Mertens, D.M. & Sass-Lehrer, M.A. 2003. *Parents and Their Deaf Children: The Early Years*. Washington DC: Gallaudet University Press.

Mitchell, R.E. & Karchmer, M.A. 2002. *Chasing the mythical ten percent: Parental hearing status of deaf and hard of hearing students in the United States*. Washington DC: Gallaudet Research Institute, Gallaudet University.

Moeller, P.M. 2000. Early intervention and language development in children who are deaf and hard of hearing. *Paediatrics*. <http://pediatrics.aappublications.org/content/106/3/e43.short>

Most, T., Ingber, S. & Heled-Ariam, E. 2011. Social competence, sense of loneliness, and speech intelligibility of young children with hearing loss in individual inclusion and group inclusion. *Deaf Students and Deaf Education* 17(2): 259–272. doi:10.1093/deafed/enr049

Nelde, P.H. 2002. Language contact. In *The Oxford Handbook of Applied Linguistics*, R.B. Kaplan (ed.), 325–334. Oxford: OUP.

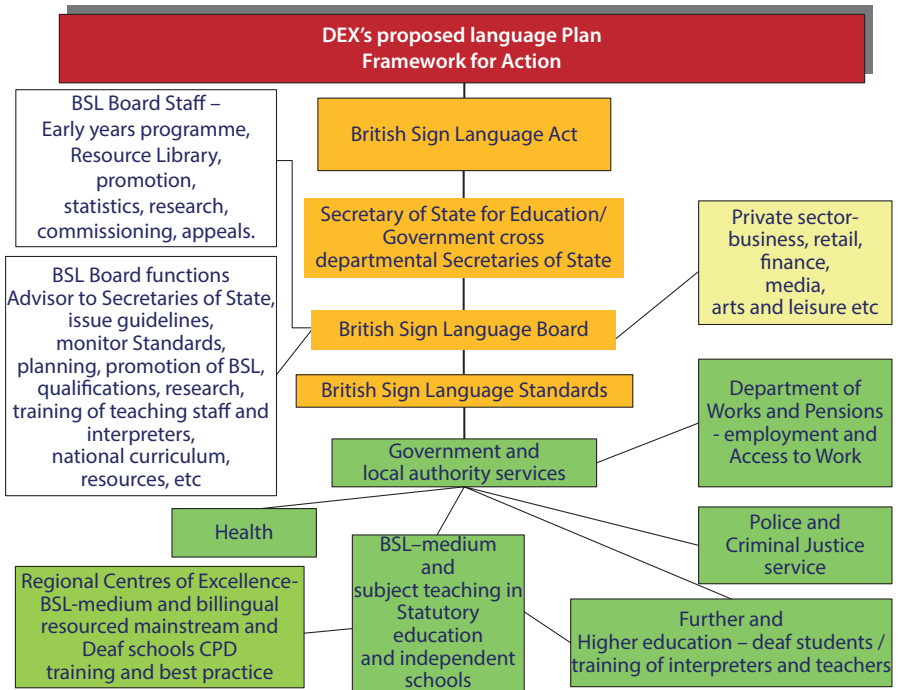New Zealand Office for Disability Issues. 2011. Review of New Zealand Sign Language Act report.

Nicolaisen, O., Trask, I., Skutnabb-Kangas, T. & Dunbar, R. 2005. Indigenous children's education and indigenous languages. Expert paper 1 written for the United Nations Permanent Forum on Indigenous Issues. <http://arcticlanguages.com/papers/PFII_indigenous_childrens_education.pdf>

Oyler, R.F. & Gross, S.R. 2000. Survey of educational preparation in paediatric audiology: A decade later. *Communication Disorders Quarterly* 21: 195–209. doi:10.1177/152574010002100402

Richardson, J.T.E., Marschark, M., Thomastine, S. & Sapere, P. 2010. Deaf and Hard-of-Hearing Students' Experiences in Mainstream and Separate Postsecondary Education. *Journal of Deaf Studies and Deaf Education* 15(4): 358–382.

Salamanca Statement and Framework for Action on Special Needs Education, 1994. <http://unesdoc.unesco.org/images/0009/000984/098427eo.pdf>

Sarant, J. & Garrard, P. 2014. Parenting stress in parents of children with cochlear implants: Relationships among parent stress, child language, and unilateral versus bilateral implants. *Deaf Students and Deaf Education* 19(1): 85–106. doi:10.1093/deafed/ent032

Schein, J.D. 1996. The demography of deafness. In *Understanding Deafness Socially*, P.C. Higgins & J.E. Nash (eds), 21–43. Springfield OH: Charles C. Thomas.

Scottish parliament's national plan for Gaelic. <http://www.scotland.gov.uk/Publications/2010/07/06161418/5> 13 October 2014.

Shirin, A., Jones, P., Luckner, J., Kreimeyer, K.H. & Reed, S. 2011. Social outcomes of students who are deaf and hard of hearing in general education classrooms. *Academic Journal Article. Exceptional Children* 77(4): 487–502.

Skutnabb-Kangas, T. 1995. Linguistic human rights, past and present. In *Linguistic Human Rights: Overcoming Linguistic Discrimination*, T. Skutnab-Kangas & R. Philipson (eds), 71–110. Berlin: Mouton de Gruyer. doi:10.1515/9783110866391

Skutnabb-Kangas, T. 2000. Linguistic genocide in education – Or worldwide diversity and human rights? In *English in International Deaf Communication* [Linguistic Insights 72], C. Kellett Bidoli & E. Ochse (eds), 75–94. Bern: Peter Lang.

Skutnabb-Kangas, T. 2003. *Language rights in a globalised world*. Montreal, Canada: Plenary address at the World Federation of Deaf People conference.

Skutnabb-Kangas, T. & Aikio-Puoskari, U. 2003. Exclusion or inclusion – Linguistic human rights for a linguistic minority, the Deaf Sign language users, and an indigenous people, the Saami. In *Many Voices, One Vision: The Right to Communicate in Practice*, P. Lee (ed.), 59–88. London: WACC

Skutnabb-Kangas, T. 2008. Bilingual education and Sign language as the mother tongue of Deaf children. In *English in International Deaf Communication* [Linguistic Insights 72], C. Kellett Bidoli & E. Ochse (eds), 75–94. Bern: Peter Lang.

Strubell, M. 2001. Edit. In *Reversing Language Shift: Theory and Practice of Assistance to Threatened Languages*, J.A. Fishman (ed.), 260–283. Clevedon: Multilingual Matters.

Stokoe, W.C. 1960. *Sign language structure: An outline of the visual communication systems of the American deaf, studies in linguistics: Occasional papers*. Buffalo NY: Department of Anthropology and Linguistics, University of Buffalo.

Svartholm, K. 2006. Review of FinSSL – Finland Swedish sign language. Nordica, Department of Scandinavian Languages and Literature, University of Helsinki. *Language and Style: Magazine for Swedish Language Research* 16: 219–223.

Swedish National Association for Deaf, Hearing-Impaired and Language Impaired Children. <http://www.dhb.se/?id=764>

The Council of Europe. *Status of sign languages in Europe report in co-operation with the Committee on the Rehabilitation and Integration of People with disabilities*. Edited by Council of Europe Publishing F-67075 Strasbourg Cedex. <http://book.coe.int>

Turner, G.H., 2009. Sign language planning: Pragmatism, pessimism and principles. *Current Issues in Language Planning* 10(3): 243–254. doi:10.1080/14664200903162505

Van Naarden, K., Decouflé, P. & Caldwell, K. 1999. Prevalence and characteristics of children with serious hearing impairment in metropolitan Atlanta, 1991–1993. *Paediatrics* 103: 570–575.

Vostanis, M., Hayes, M., Feu, D. & Warren, J. 1997. Detection of behavioural and emotional problems in deaf children and adolescents: Comparison of two rating scales. *Child: Care, Health and Development* 23(3): 233–246. doi:10.1111/j.1365-2214.1997.tb00966.x

UNESCO. 2003. Language Vitality and Endangerment. Document adopted by the International Expert Meeting on UNESCO Programme Safeguarding of Endangered Languages. <http://unesdoc.unesco.org/images/0018/001836/183699E.pdf>

UNESCO. 2009. Atlas of the World's Languages in Danger.

Wheatley, M. & Pabsch, A. 2012. *Sign Language Legislation in the European Union*. Brussels: European Union of the Deaf.

Welsh Assembly. 2014. The Welsh Ministers' response to the Welsh Language Commissioner's standards reports and advice note. <http://wales.gov.uk/docs/dcells/publications/140724-welsh-ministers-response-en.pdf>

Xie, Y.H., Potměšil, M. & Peters, B. 2014. Children who are deaf or hard of hearing in inclusive educational settings: A literature review on interactions with peers. *Journal of Deaf Studies and Deaf Education* 19(4): 423–437. doi:10.1093/deafed/enu017

Yoshinaga-Itano, C., Sedey, A., Coulter, D.K. & Mehl, A.L. 1998. Language of early and later identified children with hearing loss. *Pediatrics* 103: 1161–1171. doi:10.1542/peds.102.5.1161

Websites

Gregory, Susan. Date unknown. *Schools for the Deaf*. <http://www.batod.org.uk/index.php?id=/resources/history/histspsch/histspsch.htm>

Appendix

DEX's proposed language Plan Framework for Action

British Sign Language Act

BSL Board Staff – Early years programme, Resource Library, promotion, statistics, research, commissioning, appeals.

Secretary of State for Education/ Government cross departmental Secretaries of State

Private sector- business, retail, finance, media, arts and leisure etc

British Sign Language Board

BSL Board functions Advisor to Secretaries of State, issue guidelines, monitor Standards, planning, promotion of BSL, qualifications, research, training of teaching staff and interpreters, national curriculum, resources, etc

British Sign Language Standards

Department of Works and Pensions - employment and Access to Work

Government and local authority services

Health

Police and Criminal Justice service

Regional Centres of Excellence- BSL-medium and billingual resourced mainstream and Deaf schools CPD training and best practice

BSL–medium and subject teaching in Statutory education and independent schools

Further and Higher education – deaf students / training of interpreters and teachers

A cost-and-benefit approach to language loss*

Salikoko S. Mufwene
University of Chicago

The linguistics discourse on language endangerment and loss has been marked by a number of disputable assumptions about what languages are and about the terrible price humanity incurs in losing linguistic and cultural diversity as some of them die. I dispute some of those assumptions, including the claim that there are language rights. I also raise issues about the notions HERITAGE and ANCESTRAL languages, which should not be confused with MOTHER TONGUE. I argue that language loss is a consequence of the communicative habits of speakers, influenced in the here and now by their particular socioeconomic ecologies. The notion of POPULATION STRUCTURE, which has to do with whether a population is integrated or segregated, who gets to interact regularly with whom, and who has to accommodate whom linguistically, plays an important role in my arguments.

Keywords: adaptation, benefit, colonization, ecology, globalization, heritage language, language rights, population structure

1. Introduction

Typically driven more by a righting rather than an explanatory spirit, the linguistics discourse on language endangerment and loss has been marked by a number of disputable assumptions about what languages are and about the terrible price humanity incurs in losing linguistic and cultural diversity as some of them die. To be sure, this characterization is more typical of language advocates (e.g., Nettle & Romaine 2002; Skutnabb-Kangas 2000; Mühlhäusler 1996, 2003; Crystal 2004) than of those who have engaged in language documentation (as in Austin & Sallabank 2015). Below, I dispute some of those assumptions, including the

claim that there are language rights putatively distinct from human rights and the implicit message that any attempt to save a language really serves the needs of the population associated with the language in question (which is more present in the literature on language revitalization).

Toward the end of this essay, I also problematize notions such as HERITAGE and ANCESTRAL languages as they intersect with that of MOTHER TONGUE. This raises the question of whether an individual or a population is expected to be wedded to their heritage or ancestral culture in the same way they are to a particular race or ethnicity, as socially defined as these constructs are. These considerations should help us understand how and why languages die and whether the loss of a language necessarily makes the relevant population of speakers maladaptive or less competitive in the face of ever-changing socioeconomic ecologies. They prompt us to identify what particular ecological factors roll the dice on the vitality of particular languages and what particular actions it would take to revitalize endangered languages if linguists had the power to do so in the first place.[1]

2. Some disputable assumptions about language endangerment and loss

2.1 The cost of decreasing or losing linguistic diversity in the world

A common argument articulated by language advocates is that loss of linguistic diversity impoverishes our collective knowledge of the world, because it deprives us of one particular way in which the relevant language categorizes the world of experience and/or representation or how its speakers package chunks of information about it and/or about their experiences. Languages are thus seen primarily as representation systems (e.g., Mühlhäusler 2003 and other self-professed "ecolinguists") rather than as technologies for communication.[2] They consider the

---

1. Useful references on the subject matter of revitalization include Fishman (1991) and Grenoble & Whaley (2006). As a general policy in this essay, which is a broad critique of the literature on language loss and generally negative about most of it, I will refrain from singling out particular authors while ignoring a host of others that assume similar positions. I will thus be meager on negative citations and will identify specific authors only when there is something that justifies singling them out. I happen to agree with some of the courses of action proposed by the above cited authors, although I doubt that those initiatives alone will revitalize moribund languages or just succeed in producing knowledges that particular individuals can boast about but do not use practically in their daily lives.

2. I argue in Mufwene (2013a, in press) that the knowledge-representation aspect of language is a consequence of how a particular population has shaped their communicative technology, especially how the relevant speakers have chosen to package information in particular chunks

particular ways in which different languages categorize the world, or break it up into jigsaw-puzzle-style categories (especially through their vocabularies but also through their grammatical categories), as immutable. Languages are thus considered as repositories of several generations of accumulated knowledge, which the populations associated with them would lose if they died. Languages putatively serve as some sort of external collective memories (comparable to external drives of computers, Logan 2007) for the collective knowledges of the relevant populations. Although this population-specific knowledge is not wholly accessible to any speaker of the relevant language, it is presumably well stored in the language for any speaker to discover at some point. Faithful to the spirit of museums, languages encapsulate the histories of the populations they are associated with and should be maintained.

True to the conception of language by Ferdinand de Saussure (1916) as a social institution to which speakers are born and which they must pass on intact to later generations, the discourse sounds as if languages were anterior to the populations speaking them. Structural change is treated as the accumulation of accidental mis-productions by some speakers (showing how performance may forge divergent competence) or the consequence of materials or structures introduced from other languages. Little note has been taken of the fact that languages are actually produced and reshaped several times over by their speakers during their speech acts and in response to their current communicative needs (Mufwene 2001). Some needs die (and with them some ways of expressing ideas or feelings) and new ones arise, while some others just change, as a population's communicative ecology evolves. The latter may change when the relevant population comes in contact with another and exchanges experiences and knowledge with the other, which sometimes trigger borrowing terms or phrases that come along with the new experience or knowledge (what Bloomfield 1933 identifies as "cultural borrowings").

In some cases, it is the other population's language that is adopted. In doing this, members of the relevant population do not necessarily intend to give up their heritage language, which they have traditionally spoken. Rather, the changing ecology of their economic activities and/or social interactions makes it necessary for them to learn the language associated with the new socio-economic world order and spoken by their more influential new neighbors. Such practice of communicating increasingly in the other population's language is an adaptive response to

---

corresponding to words and phrases during their communicative acts. Knowledge can be acquired and stored without a language, but the latter is needed for those who bear it to share it explicitly with each other. Evolutionarily, this practice helps knowledge grow faster in a population.

an important ecological pressure that the individuals experience. Whether or not the changing ecology ultimately leads to the loss of the language that they have traditionally spoken, by a process of atrophy or attrition (caused by gradual and cumulative language shift, affecting the overall population), is another story to which I return below.

I submit that languages change and sometimes are replaced by others as part of their (potential) speakers' adaptive responses to pressures from the relevant communicative ecologies. Although languages also function as social identity markers and embody cross-cultural variation, they are really like other cultural tools, shaped and adapted by those who need and use them. As I explain below, knowledge of them may become rusty, when members of the relevant populations have not used them for a while, which is what language atrophy or attrition is. The experience is indeed comparable to a metallic tool becoming rusty and potentially less useful.

Languages vary as representational systems by virtue of how differently their respective speakers shape them, just like different populations manufacture technologies that are functionally equivalent in different ways (Mufwene 2013a, in press). Thus, for instance, different populations do not dress, cook or process foods, build and decorate their dwellings, organize their social groups, or practice their religions identically, even when they experience similar ecological constraints. When populations come in contact, they are likely to influence each other (though not in identical ways), with some often giving up their traditional ways and adopting those of the other populations. This is how several populations have, for instance, become Christian, Muslim, or Buddhist; have aspired at democracy, and dress more and more in the Western style, thus adopting cultural practices that are not part of their ancestral heritages. The point is that, as cultural artifacts, languages have been subject to similar adaptation pressures in human history. Despite our love for museum traditions, language advocates should make a more convincing case for saving particular endangered languages, bearing in mind that contact situations and their impacts on the relevant populations vary one from the other.

2.2 Does losing a language entail losing the culture associated with it?

A concomitant of the above is the claim that losing one's language is losing one's culture, which is allegedly disadvantageous to the individuals or population undergoing the process. How to react to this claim depends largely on how one interprets the relationship between LANGUAGE and CULTURE. I conceive of the latter notion as the particular ways members of a population behave (towards each other

and with the non-human world), how they do things, what particular beliefs they hold, and what assumptions underlie their behaviors. If one assumes that culture is separate from language, then one must consider Sapir's (1921) observation that the two do not necessarily go hand in hand. Thus, for instance, Portuguese has prevailed as the dominant vernacular of Brazil, at least among the non- or less indigenous populations (those that are not Native Americans or are of mixed Native and non-Native descent), but Brazilian culture is different from that of Portugal, despite being impregnated in part by a Portuguese cultural legacy. The same may be said not only of polities such as the United States, Canada, and Australia but also of Romance countries in Europe where the indigenous Celts first shifted to Vulgar Latin, before this language evolved into local neo-Latin varieties, which would ultimately become the modern Romance languages.

In places where a European colonial language has prevailed as the or an official language, European cultural influence is typically minimal and often superficial, as speakers of the language have typically retained most of their cultural traditions, even though these have also changed over time. And, because the larger segments of the relevant populations bear some cultural influence from Europe, such as in dressing conventions and long-distance transportation, but do not speak the European official language inherited from the colonial rule, it is evident that language and culture do not go hand in hand.

However, if one assumes that language is one of the several aspects of culture and constitutes the particular technology used by a population to communicate within itself (Mufwene, in press), then the historical evidence generally shows that those who have shifted languages have typically also indigenized the language they have appropriated (Mufwene 2009). The literature on the indigenization of English in former British exploitation colonies (typically in Africa and Asia) speaks loud and clear about structural, semantic, and pragmatic differences between the so-called "native" and "indigenized Englishes." A closer look at the "native Englishes" spoken in former English settlement colonies also reveals an important amount of indigenization interpreted as adaptation to the communicative habits of the new populations of users (some of whom had shifted from different heritage languages) and to the new physical and non-physical ecologies in which they have evolved (Mufwene 2009). This explains why we can now speak of American, Canadian, and Australian Englishes, among others, and can even identify dialects thereof with the respective polities.

Then arises the question of whether cultures are meant to be static or adapt naturally to the changing physical and social ecologies of their practitioners, who actually shape them. A correlate is whether the populations associated with particular cultures are wedded to the latter in the way they are permanently associated with particular races or ethnicities, or whether they can adapt by adopting the

culture of another population they have been in contact with. As noted above, the history of mankind provides numerous examples of cultural shifts. For instance, the Frankish population invaded Gaul but adopted the Roman style of administration that the Roman Empire had bequeathed to the region and ultimately shifted to Latin as their vernacular (as this was evolving into the Romance languages). The Romans themselves had borrowed democracy and apparently part of their mythology from the Ancient Greeks, before their leadership Christianized and helped spread Christianity to the Roman Empire. Europeans learned mathematics and much of the modern banking system (such as minting coins and usage of checks) from North African muslims during the Middle Ages.

The number of examples can be multiplied. What is evident is that sometimes a culture was adopted with its language and at some other times only it was adopted. When a culture was adopted, it has usually also been indigenized and thus transformed into a new culture. For instance, Christianity in Europe has been impregnated with some indigenous Pagan practices, just like Catholicism has indigenized in South America, in the Philippines, and in Africa and is not practiced exactly like in Europe. We cannot ignore the emergence of "syncretisms" such as Voodoo in Haiti and Candomblé in Brazil, both being blends of Catholicism and African religious traditions during the colonial period.

The bottom line is that populations are culturally adaptive, reshaping their traditions in response to changing social and physical ecological pressures. The extent of divergence from particular traditions varies from one ecology to the other. Language shift and culture shift may thus be considered normal paths of adaptation in human cultural evolution. What is worth noting is that nobody plans on giving up their traditions, nor are most of the people who participate in the process aware of what is happening. The shift is noticed in hindsight, typically too late to do anything about it. In many cases, members of the relevant population intend to be bilingual or bicultural, but they evolve in ecologies where bilingualism and biculturalism are not sustainable simply because there is increasing shortage of occasions to practice the traditional language or culture.

It is of course legitimate to stand against such evolutions, in the same way that campaigns have been undertaken to eradicate epidemics, which are deleterious to humans, and to protect our physical ecologies when they undergo changes that endanger our future. If we think that language and culture endangerment among some populations are comparable to epidemics, the question is what can be done to protect them without making the relevant populations maladaptive to their changing socioeconomic ecologies or creating more problems in the polities in which they evolve?

2.3   Were there (many) more languages spoken at the dawn of modern civilizations?

It has often been claimed that the number of languages spoken in the past, before the colonization of the rest of the world by Europe since the 15th century was much larger than today; and languages have been dying at an unprecedented fast rate. Linguists have not addressed the question of the disconnect between, on the one hand, the cladograms of genetic linguistics, which illustrate speciation (therefore the emergence of new language varieties) and the concurrent increase in the number of languages, and, on the other hand, the claims of continuous decrease of number of languages over the past half millennium.[3] I am not denying the fact that numerous languages have died as a consequence of European colonization, which is the trigger of language contact and competition. I want to direct attention to the fact that similar events have occurred several times in human history.

For instance, a concomitant of geographical and demographic expansion of Latin and of the subsequent emergence of the Romance languages in southwestern Europe is the loss of an unknown number of Celtic languages. A consequence of the colonization of England by some Germanic tribes in the 5th century has been not only the emergence of Old English by the 7th century but also the subsequent gradual loss of most of the Celtic languages of the British Isles, notwithstanding that of the handful of Germanic languages that were imported along with that of the Angles and competed with it.[4] The territorial expansion of the Arabs in North Africa and parts of the Middle East during the 7th century has spread Arabic as a vernacular at the expense of more indigenous languages survived today by Amazigh and Tuareg, for instance. At the same time, distinct national varieties of Arabic have emerged in North Africa that reflect contact with the more indigenous languages. Something similar seems to have happened with

---

3.   Note also that generally these claims have not factored in the speciation of European colonial languages, especially with the emergence of creoles and pidgins, though their numbers are certainly smaller than those of the dead and endangered languages. What I want to show below is that the evolution of language vitality has been differential and we need a more accurate picture.

4.   The names *England* and *English*, modern forms of Old English words closer to *Angle*, suggest that the language of the Angles prevailed at the expense of the competing Germanic languages, though it must have been influenced by them. Some may prefer to characterize the process as a form of koinéization. This process anticipated what can be remarked about European languages in European settlement colonies of especially the Americas and Australia, where only one colonial language has prevailed as a vernacular and the others have disappeared or are buying time, as in the case of French in Louisiana and perhaps also in Quebec (Chaudenson 2008).

the expansion of Chinese Empire, especially during the Ming and Han dynasties. Though Mandarin has not prevailed yet as the universal vernacular of the People's Republic of China, it is its dominant lingua franca and that of the Chinese Diaspora. Also, while there are still demographically important minorities speaking their heritage languages, these groups have experienced relentless ecological pressures to shift to Mandarin (also known as *Potunghua* 'common language', *Guóyǔ* 'national language', or *Huáyǔ* 'Chinese language').

All the above facts, a subset of undoubtedly many more around the world, suggest that the endangerment and loss of several indigenous languages as a consequence of the last leg of the Indo-European expansion during the past half-millennium (Mufwene 2005, 2008) is a repetition of many similar events in the history of mankind.

What I question about the assertions paraphrased above regards the part about the number of languages in the world getting smaller and smaller compared to, say, Antiquity. The claim is related to the question of how many languages were spoken by mankind at the time of the exodus from East Africa or at the beginning of agriculture about 10 kya. If we focus on the time of the dispersal out of East Africa, the question depends in part on whether one assumes monogenesis or polygenesis about the emergence of language(s) in mankind. We may also stipulate that the question is perhaps a moot one, because human populations may have already diversified linguistically and ethnically by that time. However, could there have been more diverse populations then than there are today? Isn't linguistic diversity today partly a consequence of layers of mutual colonizations, on the fission-and-fusion model, since the exodus out of East Africa (Mufwene 2013b)?

We can also assume that agriculture led to the emergence of cities as contact settings. However, did the emergent cities necessarily lead to the obliteration of languages in the rural areas? Well, precisely this evolution seems to have happened in the case of the Romance languages, as various neo-Latin vernaculars replaced the Celtic languages before they competed among themselves and some, such as Isle-de-France French (in present-day Paris) displaced their competitors. Ultimately, perhaps the number of Celtic languages displaced was greater than the number of Romance vernaculars that evolved from Latin. It would be informative to find out before language advocates paint too grim a picture of what is going on in the world today.

It appears that what we observe today partly repeats history. The number of languages displaced by English and Spanish in particular in the former settlement colonies of England and Spain respectively appears to exceed that of new vernaculars that have evolved from the colonial varieties of these languages, including English creoles. On the other hand, there are also a few other varieties that are

emerging without displacing some indigenous languages, especially indigenized varieties of English, French, and Portuguese in Africa and Asia.[5] What we really need is a balance sheet of gains and losses in the language history of the world. We will address below the question of whether language losses have typically disadvantaged the relevant populations in specific territories.

2.4   Are the languages spoken by small populations necessarily endangered today?

Another common claim is that languages spoken by small populations are more endangered than those spoken by larger ones (e.g., Mühlhäusler 1996; Crystal 2000). The claim is certainly true in the Western world, where urbanization is extensive and the dominant vernacular is also the language of the economy.[6] However, the claim has not taken into account the population structures of parts of the world where small ethnolinguistic groups (not necessarily minorities!) are endemic, where geographical mobility remains difficult, and where communication with economically and/or politically powerful outsiders remains limited, where no particular group aspires at being more important politically or economically than the others, as in sub-Saharan Africa and the mountainous interior of Papua New Guinea. Such populations sustain themselves on the practice of traditional subsistence economy, especially as they remain on the margins of world-wide globalizing economic trends, from which they experience negligible pressures for change.

   The claim of the universal natural weakness of "small languages" in relation to the powerful, expanding ones is based on a Western view of the world with locally globalized modern economies which tend to function in one dominant language and where economic success tears down traditional ethnic barriers and spreads the language of the economic system.[7] It overlooks the fractionalization of many

---

5.   We should not ignore Tryon's (2006) observation that large proportions of Pacific islanders migrating to the city or joining foreign diasporas are shifting to their national expanded pidgins or to the European colonial language, especially in the metropoles, as their vernaculars. Equally noteworthy are the reasons, also articulated by Tryon, why they have been shifting languages.

6.   This process must be understood not only as 'population shift from rural to urban areas' (Wikipedia, 25 November 2014) but also as the spread of urban living conditions to rural areas. Minorities are quickly being absorbed by both trends.

7.   This is especially evident in Anglophone North America and in Australia, where European languages other than English either have "vanished" (to borrow a term from Nettle & Romaine 2002) or are on the verge of dying, for instance, in the latter case, French in the State of Louisiana. The success of the Anglo economic system and racial integration (departing from

parts of the world that keeps several small populations autonomous and isolated from each other, especially where the expanding colonial languages function as elite lingua francas of a small minority rather than as a vernacular.

Nor does the claim take into account the fact that several economically weaker parts of the world survive on informal rather than formal economy. The more significant language in this economic practice is typically a local vernacular or a regional indigenous lingua franca (Vigouroux 2013). Their ethnographic functions are not in competition with those of either the larger ethnic groups, which are hardly interested in dominating them, or the European lingua francas spoken by the small indigenous socioeconomic elite. In this general picture, vernaculars compete with other vernaculars, when they are in regular contact, and lingua francas with other lingua francas (Mufwene 2008), though, as observed by Hagège (2014), there is a chance that a lingua franca may evolve to become a vernacular. This is an evolution that is at the mercy of how the economies of numerous Third World countries will develop, hopefully better than today.

On the other hand, one can identify situations where a major language, spoken by millions of people, may be at risk because its speakers are invested in speaking fluently a major competitor language. This may be the case for Afrikaans in South Africa, where urban affluent Colored people, who had adopted this new indigenous African language as a vernacular but owe no ethnic loyalty to the Afrikaner population, are increasingly raising children with English as the only or dominant vernacular. As more and more urban Afrikaners themselves raise English-speaking children, they may be engaging themselves in the trend of weakening the vitality of their ethnic language to the competition of English (Broeder et al. 2002), which is now rising as a vernacular among affluent urban South Africans of any ethnic background. The trend is comparable to that which spelled the loss of several European languages to the economically and politically dominant one in, for instance, Anglophone North America, Australia, and Brazil, with Portuguese prevailing in the latter case.

The reason why the vitality of Afrikaans may be eroding now is probably to be associated with the following reasons: (1) the stigma of the Apartheid regime (1948–1994) against which non-White South Africans had fought; (2) the change of political regime which in 1994 put a predominantly Black group (through the African National Congress) in political power; (3) the implicit adoption of English as the dominant official and the only international language in wide currency;

---

the earlier nationally segregated population structure of European colonists and immigrants) had the consequence of eroding the vitality of the European national languages (Mufwene 2009). Cajun French in Louisiana may be compared to Pennsylvania Dutch, both protected by the social isolation from the mainstream socioeconomic structures.

and (4) the association of the economy's white collar sector with English only. Equally important, from the point of view of world-wide economic globalization, is the selection of English by foreign investors as the work language, notwithstanding the total large size of the Anglophone countries as the primary destinations of those who consider a better political and economic climate outside the country. As much as they want to protect their cultural heritage, Afrikaners too want to maintain the economic advantages they had enjoyed during the Apartheid regime. They cannot secure them without being competitive in English. As a matter of fact, they, more than members of other ethnic groups of non-British descent, may feel more ecological pressure to be fluent in English, in urban centers, in order to remain competitive linguistically.

What linguists working on language endangerment have typically not articulated explicitly regarding language loss in especially the Anglophone North America is the fact that one must be integrated either as a tool or as a competitor in the current socioeconomic world order to experience the pressure to shift languages (Mufwene 2008, 2009). Although Native Americans lost many of their languages owing to the extermination of masses of them by Old World ills (Crosby 1992) and by especially the genocides committed by the Spaniard conquistadors, the first people to shift to the dominant European languages were the African slaves and European indentured servants that immigrated from European nations other than the dominant colonizer. These were the human tools of the colonists' economic success and hardly had the choice of keeping their heritage languages as vernaculars.

In Anglophone North America, most other European colonists lived in their national enclaves (sorts of mini-colonies within the English colonies), functioned in their national languages, and did not feel the pressure to shift to English until after the American Revolution, especially during the 19th century, after the Anglo economic system (aided by the Anglo political dominance) prevailed over the competing economic system (Mufwene 2009). It is also during this period that Native Americans started losing more of their languages, as they were losing more and more land to the expanding European invaders in the 19th century and started getting mixed on the reservations (Banner 2005). It is after this second wave of invasion that, in a third wave of disaster, Native Americans, unable to sustain themselves economically on the barren reservations, had to jump on the bandwagon of the European invaders and experience the socioeconomic pressure that African slaves and non-English Europeans have undergone before them (Mufwene & Vigouroux 2008). They saw the advantage of speaking English in order to be competitive in the new socio-economic world order.

Although we cannot deny that demographics played an important role in this process (after all, the English colonists were the majority in the wake of the

American Revolution and had chosen the term *American*, the counterpart of *Creole*, to distinguish themselves from the English in the British Isles), history had to wait till the gradual socioeconomic integration of White America to see English spread as the dominant vernacular of the United States and Canada. Other ethnic groups not associated with slavery or indentured servitude just followed suit, including Native Americans, on whom the literature on language endangerment has chosen to focus exclusively, as if their experience were unique to them.

2.5   Are less prestigious languages being replaced by more prestigious ones everywhere?

Prestige, often associated with pride, has often been invoked as a cause of language shift. The usual claim is that speakers of less prestigious languages tend to give them up for the (more) prestigious ones, or that indigenous peoples shift typically to the vernaculars of the invaders or colonists because they have lost pride in their cultural, hence linguistic heritages.[8]

An issue about this claim is how to interpret PRESTIGE, especially whether it is derived from the political, economic, or military status, or even the level of education of its speakers, bearing in mind that these factors are not mutually exclusive. Noteworthy is also the fact that they do not draw the same kind of esteem from (different) members of a population, as the less prestigious members of a population may hate everything associated with the more prestigious group. Nonetheless, here are some relevant questions: Are prestigious languages necessarily those associated with the elite segments of a population or those associated with some power, political, economic or military in particular? And when it comes to the elite class, are the intellectual elite, for instance, coextensive with the political elite? Do all powerful politicians or administrators, for instance, constitute an elite segment of a population, say, from the point of view of level of education? For some people a language (variety) is prestigious simply because it is written and functions as the standard and/or official language (variety) in a polity, regardless of whether or not it is also used in the economic sector or in the military. What may matter the most is that it is has been chosen for the school system, which sustains and spreads it. Yet for many people around the world, the language of the school system is just that, for that particular domain; it is of no use for them, because they are precluded from the socioeconomic domains in which it would be

---

8.   To cite a more recent reference, this is one of the reasons cited by at least three contributors to *The Cambridge Handbook of Endangered Languages* (2011), edited by Peter Austin and Julia Sallabank.

required or expected. Many are people around the world who have accepted the service of interpreters in the courts of law and offices of the public administration and do not mind this state of affairs.

History provides a host of counterexamples to the alleged preference for the prestigious language (variety), though in many cases one may also argue that it is a question of conflict of kinds of prestige.[9] For instance, the Franks invaded Gaul in the fifth century but shifted to Latin and the Roman-style administration the latter had inherited from Roman colonization. (Interestingly, this adaptation is contrary to what the Germanic invaders of England did, as one of their languages, Anglian, evolved into Old English, the ancestor of today's foremost world-wide lingua franca.) Both the Scandinavians and the Norman French ruled England respectively from the late 8th to the early 11th centuries and from the 11th to 13th centuries but both conquerors eventually shifted to the less prestigious indigenous language, English, though they influenced it.

Also, of the two varieties of Latin that were exported outside Rome, it is Vulgar Latin, the less prestigious variety, rather than the more prestigious Classic Latin, which evolved into the Romance languages. Neither of these Roman language varieties displaced the indigenous languages in England and in the Eastern Roman Empire, although the latter even lasted one thousand years longer than the Western Roman Empire. Even Greek (either the classical or the koiné variety), which kept Latin in its shadow in the Eastern Roman Empire, did not displace the indigenous languages. As a matter of fact, it was easily displaced by Turkish as a lingua franca during the Ottoman Empire. Likewise, Arabic spread easily as the dominant vernacular of the Middle East and North Africa but did not impose itself in Iberia even after seven centuries of colonization or prestigious coexistence with Latin in Iberia. Noteworthy is also the case of Hittite, which was also written but was replaced by languages that do not appear to have had more prestige.

The superiority of economic practices appears to account for the spread of Indo-European languages in Europe and South Asia, as well as that of Bantu languages, at the expense of Pygmy and Khoisan languages, in most of sub-Saharan Africa, although military superiority need not be ruled out in both of the latter

---

9. Sociolinguists are (more) familiar with the notion of "covert prestige" (Trudgill 1972), which accounts for why nonstandard language varieties, considered as bad by the socioeconomically more affluent members of several linguistic communities, do not necessarily lose grounds to the varieties promoted by institutions such as schools, public administration, and the white collar sector of economies. In fact, some of them thrive as important social identity markers and conveyers of public culture. The association of Jamaican Creole with Reggae music and the Rastafarian culture is a case in point, both of which have started in the grassroots of the population. Some of us also know very well how members of the working class sometimes ridicule the way members of the upper class speak.

cases. In light of the discussion in the previous sections, economic pressures (probably in combination with other factors in some cases), rather than lack of prestige on the part of the dying language, appear to account more convincingly for the present endangerment of Native American languages to the benefit of European languages.[10]

History also provides numerous examples of populations that have maintained their less or non-prestigious languages, or varieties thereof, because of their function as social identity markers and/or as indexes of particular cultural values, notwithstanding the fact that they are considered more adequate for vernacular communication.[11] In many multilingual places around the world where speakers of different languages live in separate communities but maintain egalitarian socioeconomic relations, individuals who interact with people outside their ethnic groups learn to accommodate each other in the other's language or use a regional or urban lingua franca; they find no reason to shift vernaculars. When interethnic marriages occur, relocation generally determines which language may prevail as the only or dominant vernacular in the home. If the two ethnic groups are adjacent and interact regularly with each other, nobody finds issues with being bilingual, which probably facilitated the interethnic marriage in the first place, and children are easily raised as bilinguals.

On the other hand, interethnic conflicts tend to strengthen people's loyalty to their heritage language, as this helps strengthen cooperation and solidarity within the group. There are also many cases in the Western world that speak loud and clear against the claim that speakers of stigmatized language varieties give them up for more acceptable ones, especially when the politically and/or economically dominant population is not indiscriminately assimilationist. This explains why Old Amish and African American Englishes, as well as Gullah are still spoken in the United States, as stigmatized as they are. In places like Hilton Head, South Carolina, the intrusion of wealthy property owners, which has undermined land ownership among the more traditional African American residents, has contributed to "revitalizing" Gullah, contrary to repeated claims since the late 19th century that it would die within the next generation or so (Mufwene 1997). In Jamaica, and perhaps elsewhere in Anglophone Caribbean, Creole remains strong (contrary to claims of "decreolization") because it is the identity marker of the majority population that has been marginalized socioeconomically. It is also the

---

10. See also Harbert (2010) for an elaborate discussion that complements this one.

11. Quite noteworthy is the fact that stand-up comedians often, if not typically, resort to nonstandard varieties for comical effects, or to drive a social criticism home, especially if they come from the same social background they are commenting on.

most powerful medium of expression of popular culture, especially in the local reggae and calypso music, which voice the protests of the disenfranchised.

In principle, there has always been a chance for European languages in former European exploitation colonies of Africa to become more widely used as vernaculars, just like the neo-Latin varieties did in today's European Romance countries (a shift that actually occurred after the collapse of the Western Roman Empire!). However, the socioeconomic marginalization of the majority population in all the relevant countries accounts for why in most of them the European languages are seldom used even as lingua francas outside the highly-educated elite class. The indigenous languages function as social identity markers of those that are socioeconomically marginalized; they are also the subset of indigenous languages that function as urban vernaculars in Africa and are associated with modern popular culture.

Even the elite must speak these languages outside their elite circle, for instance, in order to be able to buy goods at reasonable prices at the market, to communicate with relatives that are not as well off materially as they are, to make sure they are well understood by their domestics, and to partake in the joys of popular culture. The African elite have not entirely Europeanized, after all; for the majority of them indigenous languages remain markers of ethnic or cultural identity. The spirits of ancestors are not expected to understand foreign languages, especially when one joins the extended family in engaging in traditional rituals. Family members in the village deride their relatives that have forgotten their traditional languages and/or other cultural practices.

2.6   Have colonization and world-wide globalization uniformly caused the death of languages spoken by the economically weaker populations?

2.6.1   *What colonization does to languages of the dominated populations*
Colonization and globalization have been associated with language endangerment and loss as if they were new and uniform processes that have occurred only in the past half millennium. As noted above, the history of the dispersal of mankind out of East Africa, which continues to date (see below), has been marked by migrations and several layers of mutual colonization. The latter phenomenon is quite evident, for example, from the Hellenic Empire; its replacement and expansion (minus the loss of the Persian part) par the Roman Empire; the colonization of England by Germanic tribes; the invasion of the former Western Empire by the Franks, the Visigoths, and the Vandals; the Arab colonization of North Africa and Iberia; the replacement of the Eastern Roman Empire by the Ottoman Empire (though their borders are not coextensive); the expansion of the Chinese (especially during the

Ming and Han dynasties) into modern day China and Taiwan; and the expansion of the Bantu population. One can surely say that the European colonization of most of the world since the 15th century is a continuation, aided by innovations in modern technology (such as the invention and spread of the compass and the Caravel ship), of the Indo-European expansion, which started about 6 kya (Mufwene 2005, 2008). This process is continuing in places like the Amazon, which has brought our attention to deforestation and species endangerment. Brazilians of European descent had to wait until more efficient technology became available in the 20th century before penetrating the rain forest and beginning its economic exploitation and causing enough ecological damage in the process. The world was quickly informed about how Native Americans whose ecologies were affected suffered the consequences healthwise and culturally.

Although it is true that several cases of colonization have resulted in the replacement of indigenous languages by that of the colonizing power, the same history also shows that this has not been a uniform evolution. In some cases, as in that of the Latinization of part of Europe, the shift actually occurred after the colonizers had left, with the new trend led by a minority of indigenous or other rulers and by the new socioeconomic world order adopted by the emergent post-imperial nations.[12] As well explained by Mufwene (2001, 2008) and Osterhammel (2005), colonization has not been uniform. Its cultural consequences have been diverse, even in cases when one can claim that the same regime, such as settlement colonization, was applied. Several social and political factors influenced the process of colonization differentially. For instance, extensive genocide in the early stages of the colonization of the Americas and the Caribbean by the Spaniards contributed significantly to the numerical decrease of Native American peoples in the former and their extermination on most of the latter. Thus, genocide as a side effect of the

---

12. What has typically been misinterpreted in the case of the Romance languages is the fact that there was really no single Old French, Old Portuguese, or Old Spanish which emerged first and diversified into so many modern dialects. Instead a multitude of local or regional neo-Latin varieties emerged that would evolve into what have been survived by modern languages such as Sardinian, Sicilian, Italian, Neapolitan, Piedmontese, Venetian, Corsican, Castilian, Catalan, Galician, Provençal, Occitan, Walloon, Picard, Normand, Romansh, Madeirense, and Estremenho, among many others, which individually displaced the more indigenous Celtic languages of their regions. These neo-Latin varieties have also competed with each other, with one variety promoted over others as the national official language. This is most obvious in the case of modern French, for the protection of which the Académie Française was created and for which the whole school system was mobilized, especially in the 20th century, to make sure that no other language would be spoken within the European boundaries of the French Republic. It is also obvious in the fight of Catalonia for its political autonomy and for the recognition of Catalan as a co-official language of Spain, next to Castilian (known to most outsiders as Spanish).

colonization drove some of the indigenous languages out. Planned miscegenation in Latin America also reduced drastically the proportion of pure Native Americans at the expense of indigenous languages; the Mestizos typically Hispanicized linguistically while shaping new, Latin American cultures.

In English and French colonies, it is especially the ills from the Old World, and since the late 19th century, the further incursion into Native American reservations and rapid urbanization (in the sense explained above) which, based on the numbers in Nettle and Romaine (2000), reduced even more extensively the number of Native American languages in the United States and Canada (less drastically in the latter than in the former).[13] These differences in styles of colonization should keep us from interpreting the differential loss of indigenous languages as being the consequence of the same colonization activities everywhere.

I would be remiss not to discuss the often-cited impact of boarding schools in various colonies, where indigenous children were educated in the colonizing nation's language and often prepared to function as colonial auxiliaries. For many parts of the world, this impact should not be exaggerated. In the case of Africa and Asia, graduates from these schools became colonial intermediaries and thus interpreters, interfacing between the colonizers and the masses of the unschooled Natives. They did not necessarily stop speaking the indigenous languages, even at a later time when there were also educated women, whom the educated men

---

13. The contrast I articulate here to distinguish the beginnings of Spanish colonization from those of English and French colonization is not intended to ignore the large numbers of Native Americans that died during confrontations with White Americans during especially the westward expansion of the United States. However, as pointed out by Crosby (1992), ills did a deadlier job than fire weapons in helping Europeans conquer the Americas. Also, I do not mean to suggest that miscegenation with Native Americans did not occur at all in English and French colonies. The evidence for miscegenation is most obvious from the Michif (< French *métis* 'Mestizo'), who are genetically hybrids between the French *coureurs de bois* and Cree Indians. They produced a new, "intertwined" language whose verbal structure is Cree but the nominal one largely French. However, unlike the Latin American Mestizos, the Michifs have, to my knowledge, remained as socioeconomically marginalized as pure Native Americans to date. Notwithstanding cases of individual colonists who married Native American women, it is not clear whether there were no Michif-like communities, albeit dying ones, in English North America. In Northern Alaska, Russian colonists not only formed unions with Native Alaskans but also produced Copper Island Aleut (also known as Mednyj Aleut) with them, which is still spoken by a few of the descendants of the mixed offspring. Overall, these are marginal cases compared to the Latin American hybridization and Hispanicizing phenomena, which have reduced Native American populations to negligible demographics, for instance, 2% in Costa Rica, 17.5% Native Americans vs. 65% Mestizos in Mexico, and 0.4% vs. 43.1% Pardos in Brazil. (The Wikipedia entries of many countries prefer giving linguistic statistics to ethnic or racial ones, but these are not more reassuring, for instance with Quecha spoken only by 13% of the population in Peru).

(who had been favored all along for schooling) could marry. Interethnic marriages between educated men and women did not force them to give up their ethnic languages either, because they continued to interact with members of their respective ethnic groups. The real exceptions were when the spouses had met at a place where the language of neither of them was spoken; but several of them were also likely to have learned the local/regional indigenous lingua franca, which they may have found more natural to speak when they courted each other. The local population structure enabling this absorbed them as different from the European colonial agents, who did not socialize with nor assimilate them in the first place. Inter-ethnic marriages in Africa are also a late, 20th-century phenomenon that has been particularly urban. Children from such unions acquire the indigenous urban vernacular at least as one their mother tongues.

Regarding the relocation of indigenous people as a possible cause of language loss, the British may have accomplished a lot of this, especially in using the educated Indians in the colonization of other Asian territories east of India and transporting large numbers of contract laborers from India to Pacific and Atlantic colonies. Those who went to Trinidad and Guyana, for instance, shifted to the local English creole as their vernacular. In South Africa their descendants have become vernacular speakers of English or Afrikaans. However, as is obvious now in India, what this population relocation practice accomplished was certainly the decrease of a number of learners of heritage languages. Since the proportion of such civil servants was relatively small, the impact of the practice on the vitality of Indian languages is negligible. That of contract laborers may have to be assessed, in the same way one must wonder what the linguistic impact of the slave trade of the 16th–19th centuries was in sub-Saharan Africa, given the depopulation it must have caused in some regions.

A peculiarity of Anglophone North America, colonized on the settlement model, is that Native American children were schooled in English with the intention of deculturating them from their heritages but not necessarily to integrate them and make them economically competitive within the new, White-majority American population structure. They were taught professions that were considered suitable for the reservations. The purpose of Anglicizing the children appears to have been to facilitate communication between the Americans (predominantly of European descent) and the Native Americans, with apparently the interests of the socioeconomically dominant population in mind.

What is not clear is whether graduates from these schools all returned to the reservations and, if they did, whether they were unable to recover competence in their heritage languages to communicate with their relatives who had not been schooled. Could their relatives give up their ancestral languages to speak only English just because the school children had forgotten them? Wouldn't the

schooled children have tried to re-acculturate back by resurrecting some competence in their native languages and being bilingual? Or did this deculturation experiment exert its most negative impact on the vitality of indigenous languages only if the boarding school graduates left the reservations altogether and went to live in the city and thereby permanently adopted English as their vernacular? In the latter case, the school-in-English practice would have contributed to reducing the demographic size of speakers of Native American languages without however spreading the new vernacular among the indigenous peoples. Thus, are Native American languages not endangered primarily for reasons that have more to do with the lure of urban life and the new, European-style socioeconomic world order than with schooling in English, although this facilitated functioning in the city?

Likewise, the (alternative) hypothesis that Native American and Australian Aboriginal languages have been dying because of continuous emigration from reservations to the city brings to mind other cases of language loss associated with massive or successive exoduses from the homeland. A most notable case is that of Biblical Hebrew, a consequence of the dispersal of Jews to various places originally around the Mediterranean and then to various parts of the world, especially Europe, where some of the Diaspora populations became associated with language varieties such as Judeo Spanish and Yiddish. These varieties are consequences of language shift in the host countries where the Jews were not socially integrated, or did not assimilate, with the Natives but did not have the critical mass to maintain their ancestral language, often after several generations of relocation to other places. Later dispersals under similar conditions spread these particular language varieties elsewhere, where they became ethnic markers when the Jewish communities were segregated. Changing political atmospheres in Europe and the ultimate relocation of many of their speakers to Israel, where Israeli Hebrew (different from Biblical Hebrew) had emerged as the national language spelled the death of the Diaspora language varieties, both by demographic attrition in places where they had been spoken and by communal vernacular shift to Israeli Hebrew.

Another case worth considering is that of Irish. Is it endangered just because of the new socioeconomic conditions since especially the 19th century that have favored English as a vernacular? Or did the "Great Famine" that reduced the Irish population by the deaths of about 1 million people and the emigration of another million or so (a total of 20–25% of the population) reduce the resistance of the indigenous language to that of the colonizer? Should we ignore the massive emigration to the United States and to various places in the British Empire after Hiberno English, the national indigenized variety of English, was instituted as the medium of formal education?

Note also that the independence of the Republic of Ireland did not drive English out from the economic life of its citizens. Though most children have been

taught Irish as a subject in school for the past century or longer but have not had to use it in their adult lives, what appears to be a notional commitment does not appear to have helped the vitality of Irish significantly. The effects are thus comparable to those of the teaching of and in French in schools of Francophone Africa since the 1960s. Contrary to expectations, the proportion of actual speakers of French has not exceeded an estimated 20% of the population in most countries. An important reason is that most learners have no reason for speaking French outside the classroom.[14]

2.6.2 *Is world-wide globalization that dangerous to languages of the weaker populations?*[15]

As for the impact of globalization on the vitality of "indigenous languages," we must start by clarifying the phenomenon of GLOBALIZATION itself. As a network of interdependencies and interconnected parts in an economic system, it need not be a world-wide phenomenon. It can be observed concretely at the local level within the economic systems of cities, such as when different contractors (for example for masonry, electricity, water/plumbing, and telephone) work together toward the construction of a modern building, as well as when a city depends on an adequate infrastructure of transportation, telecommunication, and sanitation services to function adequately. Non-locally, it is evident for instance when one place depends on materials coming from another place to manufacture a product.

It is of course necessary for the different parties involved in such a global economic structure to understand each other, but not everybody need speak the same language for the system to work. Communication can be accomplished through intermediaries/interpreters, although the process may be slower. It so happens that places with a high globalization index have typically evolved toward

---

14. As noted above, it would certainly be informative to know the extent to which the African slave trade of the 16th through to the early 19th centuries (late 19th century for Brazil, Cuba, and the Dominican Republic) endangered particular languages in Africa or perhaps drove some to extinction. Did enslavement in general impact the vitality of some languages in the history of mankind? Related to this is the diaspora phenomenon of especially the Polynesian and Micronesian islands to which Tryon (2006) perceptively draws attention. Often as many people as, and sometimes more than, those remaining on the islands now live outside and lose competence in their heritage languages. At the same time, on the islands, the lure of modernity in urban centers favors the demographic expansion and further vernacularization of the pidgins which emerged during the 19th century. Tryon rightfully traces the diaspora phenomenon back to the trans-insular recruitment of contract laborers who worked on the sugarcane plantations, where the pidgins emerged.

15. Weakness here has to do primarily with the demographic size of the population. The literature has typically presented small populations as linguistically weaker than larger ones.

monolingualism, especially where almost anybody has the potential to interact with somebody else, such as in buying goods for some work one is doing or for a project one is engaged in. However, on a larger, international scale, intermediaries have interfaced companies to date. Thus, Microsoft operates in the national languages of the countries where they conduct business; and, likewise, Japanese automobile manufactures do not use Japanese to sell their cars outside Japan. So, English is far from driving world-wide globalization; on the contrary, it benefits from the economic phenomenon. To the extent that the "buyer's language" principle applies (Ostler 2005), the USA and the British Commonwealth represent a huge group of English speakers, while China represents an unbeatable population of potential learners. English is not the only major language to benefit from world-wide globalization either. Western nations doing business in China offer Mandarin courses as an option to their college students specializing in foreign languages, so that some of them may represent national companies in China or serve as interpreters in negotiations of contracts between China and them.

The dynamics of which language benefits the most depend largely on colonial history and the economic powers of the countries involved. A better understanding of the history of the world from the point of view of economic globalization (locally and non-locally) will help us question the myth that English is the "killer language" par excellence, if it had any agency at all in the process. The killers are speakers who, under particular ecological pressures, favor one particular language over others.

I won't deny that English is the most widely used lingua franca in the world today, though Mandarin and Spanish, for example, have more vernacular speakers. English just happens to have more lingua franca speakers, aided by populous countries such as India and China from the point of view of learners. This is largely a legacy of colonization that spread English as the official language of numerous exploitation colonies, including India and Nigeria, which for the purposes of world trade, count as Anglophone countries, in addition to the UK, the United States, Canada, and Australia. The more countries count as Anglophone, the more other countries find it useful to teach/learn English as an international lingua franca for the purposes of trading with the Anglophone countries. Of course it makes an important difference when some of the Anglophone countries are also the most powerful economically and militarily and, in addition, are leaders in science and technology, which has favored English over Spanish. Although the latter has more native speakers than English and is spoken in about twenty-one countries (including Spain and its former settlement colonies), it has far fewer non-native speakers (Ammon 2010). Besides, the Spanish-speaking countries are weaker economically and militarily, aside from not counting as leaders in science and technology compared to the United States, the UK, Canada, and Australia.

These factors conspire to assign Spanish a lower Q-value (de Swaan 2001, 2010) than English, making it generally less attractive as an international language or disqualifying it as a global one (Ammon 2010).

We should also remember, regarding non-local globalization, that this process is far from being recent and need not be associated with the collapse of the Soviet Union and the dismantling of the Iron Curtain that separated it from Western Europe. The colonization of the rest of the world by Europe was expanding trade networks, which were being reorganized at the end of the Middle Ages, with the Chinese in the Far East, the Arabs around the Mediterranean and in the Indian Ocean, and the Ottomans claiming the eastern Mediterranean and the Persian Gulf around the same time. The Hellenic and Roman Empires had already developed non-local global-economy networks, which help us better understand the saying "All roads lead to Rome." Even these had been preceded by the Phoenicians around the Mediterranean and the Sumerians along the Persian Gulf, among others, under the constraints of the knowledge of the world and of the technology that could be used for long-distance travel at that time. The fundamental principles remain the same, though the scope is much wider and more complex today (Chanda 2007). As explained above, the impacts on language vitality have varied from one case to another. Overall, it is not world-wide globalization that rolls the dice on the vitality of languages; it is the local population structure, subject also to the kind of local globalization index it has achieved. The higher the local globalization index, the stronger the tendency toward monolingualism.

Population structure also regards the extent of socioeconomic integration and the extent to which the dominant model, such as the Anglo economic system in the USA or the Roman-style administration and economy in Gaul and Iberia, has been assimilationist. A better understanding of POPULATION STRUCTURE will help us understand why Europe and its settlement colonies have evolved toward monolingualism, why parts of Europe are so Balkanized, why some countries risk breaking up (e.g., Spain and the United Kingdom), why European minorities are fighting for the right to operate in their heritage languages, but also why language endangerment in Africa is not only less extensive but not benefiting European colonial languages, with perhaps the exception of South Africa (as explained above).

3.   By way of conclusions

Humans are now recognized, from the point of view of evolution, to be engaged in niche-construction (Odling-Smee & Leland 2003). Although our vitality depends on the ecologies in which we evolve, we can also modify these ecologies and then undergo the consequences of the changes we have introduced. In other

words, we need not be passive patients of an evolution that is driven by changing ecologies on which we would have no control. We can influence some of our evolutionary trajectories, and in the case of languages, by modifying some relevant ecological conditions that affect them. To make sense of how our languages influence us and we affect their vitality, we must first understand how languages die and why (Mufwene 2006, 2007).

In a nutshell, languages die, gradually, as speakers practice them less and less, because they have fewer and fewer opportunities to use them, basically in the same way people forget to adequately use technology that they have not often had need for. Since knowledge of a language, as of a particular technology, resides in the mind of practitioner, a language as a communal phenomenon dies when the practices of individual speakers, in not using it, converge. For all intents and purposes a language is dead when there are no more speakers left that communicate in it, even if there are survivors who remember bits and pieces of it. One may argue that a language is not only a communicative system used by a particular population but also the reality of practicing it. It lives in the practices of its speakers not in the knowledge people have of it.

It is thus important to articulate some of the factors that prevent speakers from practicing a particular language in their community. Very simply, they evolve in settings where the relevant language, considered an alternative to another, is not needed any more or is superseded (in the simple sense of being 'pushed aside') by another. This may happen when the relevant speakers relocate to a community where they are socially integrated but another language is spoken, or when they have been dominated by another population that imposed a new socioeconomic structure in which a new language provides various advantages that the traditional language does not offer, such as with education, which opens up doors to jobs that offer better incomes and better living conditions (see also Harbert 2010). A key to this is that the dominant population must offer some sort of integration within the new socioeconomic world order, in which everybody can compete for the same or some jobs, although there is no guarantee that everybody is treated equally.[16]

There are many other reasons why languages die, but the point is that there are decreasing opportunities for speaking the language that becomes endangered. Things may start with a population structure where members of the population that is politically, economically, and/or militarily weaker use the traditional language at home and in some other private or traditional settings, but those who

---

16. As is evident from Vaillancourt (2008), a sure way of maintaining the language of a minority population is to keep the latter on the margins of and separate from the dominant population. Once its members participate in the mainstream socioeconomic structure, their language and culture face the competition of its language and culture too.

participate in some activities with the dominant population communicate with the latter in the new language. However, as the population structure changes and more and more members of the weaker population, which may or may not be indigenous, participate in the life style of the dominant population, they not only become more fluent in it but also interact with each other in it, especially in settings where they are ethnolinguistically mixed. The traditional language may become superfluous, especially when the weaker population speaks the dominant population's language even at home, because fluency in it keeps them competitive outside the home. As the older generations die, the survivors become either dominant or monolingual in the dominant population's language, although they may maintain a separate cultural identity. When they are also socially integrated, only some memory of the heritage language survives in their minds.

Whoever wants to revitalize a moribund language should realize that it is not dictionaries and grammar books that will do the job, although they will be helpful, because the knowledge encapsulated in these references is not the language itself. The language can be inferred from the communicative activities of members of the population in forms that may not even correspond to what especially grammar books present. True revitalization entails recreating a population structure in which the relevant language can be practiced, bearing in mind that a population can shift languages while simply adapting their traditional culture to the new socioeconomic structure. Even revitalizing a dying culture, if one believes that every population is wedded to its heritage culture the same way it is permanently associated with its race, entails recreating an environment in which the culture can be practiced. In other words, one should engage in the endeavor in more or less the same way an environmentalist would engage in saving or protecting an endangered species, preferably by restoring the ecology that would be advantageous to it, one in which it can reproduce itself naturally.

As is evident from the contrasting experiences of, on the one hand, French in Québec and Afrikaans in South Africa versus, and on the other hand, Irish in Ireland, it is not enough to teach a language in school to revitalize it. One should invest the socioeconomic machinery into it too. Contrary to the former two cases, a century or so of teaching Irish in the Republic of Ireland's schools or raising children in Gaeltachtaí have produced very few actual speakers of Irish. If Afrikaans may be at risk today, it is because it is losing to English some of the economic advantages, especially the most prestigious ones, which it used to provide to its speakers. It may also suffer the stigma of the exclusionary way in which the Apartheid regime disadvantaged the non-White, especially Black, segment of the South African population. The present political and socioeconomic structure may be having a revenge in disguise.

French has been revitalized in Quebec because the Québécois government engaged all economic sectors in the Province to function in French too. This commitment accounts for what may also be characterized as "the Québécois miracle," as the Québécois Francophones have exceptionally managed to maintain their language and culture (naturally not really intact!) in a wider North American socioeconomic environment north of Mexico where every (European) language other than English has either died or is moribund.

It appears that the killers of the dead and dying languages are not really the languages that displace them but the speakers who give them up or stop transmitting them to their children under the particular ecological pressures they face (Mufwene 2000). Assuming as I do that languages are communication tools, which fashion their culture-specific world views as a consequence of the particular ways in which their speakers have shaped them (Mufwene 2013a), are there any language rights worth talking about that supersede those of the people who through their communicative behaviors roll the dice on their vitality? For linguists who argue that it is possible for the speakers (who can certainly be construed as victims of the particular ecologies to which they adapt) to remain bilingual in both their heritage language and the new vernacular, what particular ecologies would sustain the proposed bilingualism? Is it deliberately that people favor the same language at home and outside home? How do many bilinguals become more dominant in one than the other? Is this situation different from that which leads to monolingualism, as among many Americans of European descent?

Speaking of heritage languages, which must be maintained (or which language advocates wish to maintain apparently under any conditions), what is the heritage language of a child: that of his/her parents or that of the social ecology in which they grow up and to which they adapt? Does one inherit one's culture in the same way that one inherits their genotypes? How many generations does one have to count back (ignoring cases of ethnolinguistically mixed unions) to feel free to ignore their ancestral heritage? Otherwise, aren't we opening a huge can of worms from a cultural evolution perspective? Note that English, claimed to be the foremost "killer language," is barely a millennium and a half old, which is a very short time relative to the dispersal of Homo sapiens out of East Africa. The people that have benefited from its dispersal around the world over the past century, or even over the past half-millennium, may not all legitimately claim it as their heritage language, if they should count many more generations back. They will discover that some of their distant ancestors actually shifted to English from some other European language. As noted above, the dispersal history of Homo sapiens, which continues to date (Chanda 2007), has been marked by layers and layers of migrations and occasional mutual colonization, language contact and competition, the displacement of some languages by others, and the emergence

of new language vernaculars or lingua francas. What one claims as their heritage culture or language may be the consequence of language shift, sometimes a recent one, which appears to have been beneficial to their parents or other ancestors.

Let me conclude by reiterating that I don't advocate letting the evolution of languages continue uniformly like in the past, especially when it is evidently adversative to some populations. However, very often, we must be content with solutions that are satisficing rather than optimal. The choice is sometimes between two evils, so to speak. For language advocates, such as those who claim that schooling children in a language other than their ancestral one is "a crime against humanity" (Skutnabb-Kangas & Dunbar 2010), the question is whether they have a solution that they can offer that is likely to work better. As many parents have feared, in some socioeconomic ecologies, schooling one's child in a language other than that of the dominant economy may disadvantage the child, because it will make him or her less competitive on the job market. If languages have any rights at all (after all they may be considered as tools at the service of their speakers!), there are indeed situations in which conflicts arise between the alleged right of a language to be maintained and the human right of a population to be competitive socioeconomically. To the concerned parents, the choice is quite clear.

As many parents who have relocated to different ethnolinguistic places (towns, cities, or countries) have realized, children can quickly overcome the disadvantage of not knowing the local language, learning it by immersion and becoming competitive in relation to their peers. For adults who relocate, the question is whether they want to adapt to the local socioeconomic structure, claim accommodations that may appear unrealistic to the indigenous population (especially if the newcomers are a small minority they can help integrate), or do violence to the local population by changing the current socioeconomic structure (as is evident from European settlement colonization of the past half-millennium).

The above remarks should not be construed as arguments against educating children in their mother tongue (which need not be the ancestral language ethnically defined), because education in one's mother tongue, in fact in the language that the child is fluent in, offers undisputable advantages to the learner. However, as in many cases the child's mother tongue is not the ancestral language, how does one reconcile the alleged right to maintain one's ancestral or heritage language with that of receiving education in one's mother tongue, especially if the professed goal of the school system is to prepare children to function adequately and to be competitive on the job market in their societies?

Linguists who argue that all living languages must be maintained for the sake of linguistic diversity have to make a more convincing case, especially when the relevant populations feel they are disadvantaged by them. Those who argue that maintaining linguistic diversity is useful to linguistics as a profession should feel

ashamed of themselves if they ignore the odds that the relevant populations are facing. The reason is professionally selfish. Well-intended linguists face ecologically what is literally a wicked problem: what is good for the survival of a particular population in the face of a changing socioeconomic ecology versus what is ideal for the practice of linguistics.

Krauss (1992) warned that, unlike ecologists (popularly identified as environmentalists) linguists may go down in history as the experts that did not care about the disappearance of the subject matter of their discipline, viz., languages. The question is whether linguists should not start by acknowledging what they can (help) change and what they cannot. This can start by acknowledging that dictionaries and grammar books are useful museum artifacts relative to the languages they describe. Although they can certainly also be helpful in efforts to revitalize languages, they will not do the job intended by those engaged in language revitalization outside socioeconomic ecologies that foster speaking the relevant languages among members of the same ethnolinguistic group, even the children, even outside home.

References

Ammon, U. 2010. World languages: Trends and futures. In Coupland (ed.), 101–122. doi:10.1002/9781444324068.ch4

Austin, P.K. & Sallabank, J. (eds). 2011. *The Cambridge Handbook of Endangered Languages*. Cambridge: CUP. doi:10.1017/CBO9780511975981

Austin, P.K. & Sallabank, J. (eds). 2015. *Endangered Languages: Beliefs and Ideologies in Language Documentation and Revitalization*. Oxford: OUP.

Banner, S. 2005. *How the Indians Lost their Land: Law and Power on the Frontier*. Cambridge: Belknab Press. doi:10.4159/9780674020535

Bloomfield, L. 1933. *Language*. New York NY: Holt, Rinehart & Winston.

Broeder, P., Extra, G. & Maartens, J. 2002. *Multilingualism in South Africa with a Focus on KwaZulu Natal and Metropolitan Durban* [Occasional Papers 7]. Cape Town: PRAESA.

Chanda, N. 2007. *Bound Together: How Traders, Preachers, Adventurers, and Warriors shaped Globalization*. New Haven CT: Yale University Press.

Chaudenson, R. 2008. On the futurology of linguistic development. In *Globalization and Language Vitality: Perspectives from Africa*, C.B. Vigouroux & S.S. Mufwene (eds), 171–190. London: Continuum.

Coupland, N. (ed.). 2010. *The Handbook of Language and Globalization*. Malden MA: Wiley-Blackwell. doi:10.1002/9781444324068

Crosby, A.W. 1992. Ills. In *Atlantic American Societies: From Columbus through Abolition 1492–1888*, A.L. Karras & J.R. McNeill (eds), 19–39. London: Routledge.

Crystal, D. 2000. *Language Death*. Cambridge: CUP. doi:10.1017/CBO9781139106856

Crystal, D. 2004. *The Language Revolution*. Malden MA: Polity.

Cunningham, D., Ingram, D.E. & Sumbuk, K. (eds). 2006. *Language Diversity in the Pacific: Endangerment and Survival*. Clevedon: Multilingual Matters.

De Swaan, A. 2001. *Words of the World: The Global Language System*. Cambridge: Polity.

De Swaan, A. 2010. Language systems. In Coupland (ed.), 56–76.
doi:10.1002/9781444324068.ch2

Fishman, J.A. 1991. *Reversing Language Shift*. Clevedon: Multilingual Matters.

Grenoble, L.A. & Whaley, L.J. 2006. *Saving Languages: An Introduction to Language Revitalization*. Cambridge: CUP.

Hagège, C. 2014. La confrontation entre l'anglais et le français dans le monde contemporain (début du XXe siècle). In *Colonisation, globalisation et vitalité du français*, S.S. Mufwene & C.B. Vigouroux (eds), 47–74. Paris: Odile Jacob.

Harbert, W. 2010. Endangered languages and economic development. In Austin & Sallabank (eds), 403–445. doi:10.1017/CBO9780511975981.020

Krauss, M.E. 1992. The world's languages in crisis. *Language* 68: 4–10.
doi:10.1353/lan.1992.0075

Logan, R.K. 2007. *The Extended Mind: The Emergence of Language, the Human Mind, and Culture*. Toronto: University of Toronto Press.

Mufwene, S.S. 1997. The ecology of Gullah's survival. *American Speech* 72: 69–83.
doi:10.2307/455608

Mufwene, S.S. 2000. Languages do not kill languages, speakers do. *The University of Chicago Magazine* 93(2): 30–31.

Mufwene, S.S. 2001. *The Ecology of Language Evolution*. Cambridge: Cambridge University Press.

Mufwene, S.S. 2005. *Créoles, écologie sociale, évolution linguistique*. Paris: L'Harmattan.

Mufwene, S.S. 2006. Language endangerment: An embarrassment for linguistics. In *The Panels: Papers from the 42nd Annual Meeting of the Chicago Linguistic Society*, J. Bunting, S. Desai, R. Peachey, C. Straughn & Z. Tomkova (eds), 111–140. Chicago IL: CLS.

Mufwene, S.S. 2007. How languages die. In *Combat pour les langues du monde – Fighting for the World's Languages: Hommage à Claude Hagège*, J. Fernandez-Vest (ed.), 377–388. Paris: L'Harmattan.

Mufwene, S.S. 2008. *Language Evolution: Contact, Competition and Change*. London: Continuum Press.

Mufwene, S.S. 2009. The indigenization of English in North America. In *World Englishes: Problems, Properties, Prospects. Selected Papers from the 13th IAWE Conference* [Varieties of English around the World G40], T. Hoffmann & L. Siebers (eds), 353–368. Amsterdam: John Benjamins. doi:10.1075/veaw.g40.21muf

Mufwene, S.S. 2013a. Language as technology: Some questions that evolutionary linguistics should address. In *In Search of Universal Grammar: From Norse to Zoque* [Linguistik Aktuell/Linguistics Today 202], T. Lohndal (ed.), 327–358. Amsterdam: John Benjamins.
doi:10.1075/la.202.22muf

Mufwene, S.S. 2013b. What African linguistics can contribute to evolutionary linguistics. In *Selected Proceedings of the 43rd Annual Conference on African Linguistics: Linguistic Interfaces in African Languages*, O.O. Orie & K. Wu (eds), 52–67. Somerville MA: Casadilla Press.

Mufwene, S.S. In press. Language evolution, by exaptation, with the mind leading. In *Essays in Evolutionary Linguistics*, G. Peng (ed.). Hong Kong: Chinese University of Hong Kong Press.

Mufwene, S.S. & Vigouroux, C.B. 2008. Colonization, globalization and language vitality in Africa: An introduction. In *Globalization and Language Vitality: Perspectives from Africa*, C.B. Vigouroux & S.S. Mufwene (eds), 1–31. London: Continuum.

Mühlhäusler, P. 1996. *Linguistic Ecology: Language Change and Linguistic Imperialism in the Pacific Region*. London: Routledge. doi:10.4324/9780203211281

Mühlhäusler, P. 2003. *Language of Environment: Environment of Language: A Course in Ecolinguistics*. London: Battlebridge.

Nettle, D. & Romaine, S. 2002. *Vanishing Languages: The Extinction of the World's Languages*. Oxford: OUP.

Odling-Smee, J.F., Laland, K.N. & Feldman, M.W. 2003. *Niche Construction: The Neglected Process in Evolution*. Princeton NJ: Princeton University Press.

Osterhammel, J. 2005. *Colonialism: A Theoretical Overview*. Princeton NJ: Markus Wiener.

Ostler, N. 2005. *Empires of the Word: A Language History of the World*. New York NY: Harper Collins.

Sapir, E. 1921. *Language: An Introduction to the Study of Speech*. New York NY: Harcourt, Brace & World.

de Saussure, F. 1916. *Cours de linguistique générale*, C. Bally & A. Sechehaye (eds) in collaboration with A. Riedlinger. Paris: Payot.

Skutnabb-Kangas, T. 2000. *Linguistic Genocide in Education, or Worldwide Diversity and Human Rights?* Mahwah NJ: Lawrence Erlbaum Associates.

Skutnabb-Kangas, T. & Dunbar, R. 2010. Indigenous children's education as linguistic genocide and a crime against humanity? A global view. *Gáldu Čála – Journal of Indigenous Peoples Rights* 1.

Trudgill, P. 1972. Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society* 1: 175–195. doi:10.1017/S0047404500000488

Tryon, D. 2006. Language endangerment and globalization in the Pacific. In Cunningham, Ingram & Sumbuk (eds), 97–111.

Vaillancourt, F. 2008. Language and poverty: Measurements, determinants, and policy responses. In *Language and Poverty*, W. Harbert (ed.), 147–160. Bristol: Multilingual Matters.

Vigouroux, C.B. 2013. Informal economy and language practice in the context of migrations. In *Language, Migration and Social Inequalities*, A. Duchêne, M. Moyer & C. Roberts (eds), 225–247. Bristol: Multilingual Matters.

Language documentation, ethno-history and language vitality

Language documentation 20 years on

Peter K. Austin

School of Oriental and African Studies (SOAS), University of London

In the last decade of the 20th century a new field of language research emerged that has come to be known as 'language documentation' or 'documentary linguistics' (Himmelmann 1998, 2002, 2006; Lehmann 2001; Austin 2010; Grenoble 2010; Woodbury 2003, 2011). In this paper we explore how it was defined in the seminal work of Himmelmann (1998) and others, including what were presented as significant characteristics that distinguished language documentation from language description, and how the field has changed and evolved over the past 20 years. A focus on best practices, standards, tools and models for documentary corpora appeared in the early years, which led later to more critical discussions of the goals and methods of language documentation. The paper examines some current developments, including new approaches to language archiving, and suggests that there are opportunities for language documentation to adopt a more socially-engaged approach to languages and linguistic research, including better engagement with language revitalisation. There are also opportunities to work towards addressing what is currently a language documentation output gap through experimentation with new genres and innovations in writing and publication.

Keywords: archiving, documentary corpora, documentary linguistics, language archiving, language documentation, language revitalisation

1. Defining language documentation[1]

Language documentation (also known by the term 'documentary linguistics') aims, according to the seminal definition in Himmelmann (1998: 161), 'to provide a comprehensive record of the linguistic practices characteristic of a given

---

speech community… This… differs fundamentally from… language description [which] aims at the record of a language… as a system of abstract elements, constructions, and rules.' Himmelmann (2006) presents it as the subfield of linguistics that is 'concerned with the methods, tools, and theoretical underpinnings for compiling a representative and lasting multipurpose record of a natural language or one of its varieties' (Himmelmann 2006: v). Language documentation is by its nature multi-disciplinary, and as Woodbury (2011) notes, it is not restricted to theory and methods from linguistics but draws on 'concepts and techniques from linguistics, ethnography, psychology, computer science, recording arts, and more' (see Harrison 2005; Coelho 2005; Eisenbeiss 2005 for arguments).

Documentary linguistics has developed over the past 20 years as a response to the growing realisation among linguists, dating from the late 1980s,[2] that a majority of the world's 7,000 languages are endangered, in the sense that they are being spoken by decreasing and aging populations in reducing numbers of domains and are not being passed on to the next generation of speakers (Robins & Uhlenbeck 1991; Hale et al. 1992; Crystal 2000; Austin 2007; Whalen 2004; Grenoble 2011). A desire among some researchers to create a lasting, and potentially unrepeatable, record of language use in its social and cultural context was one of the driving forces behind the interest in this new approach. This involved a renewed attention to context, influenced by the ethnography of communication (pioneered by Hymes 1964), and the discourse-based approach of Sherzer (1987).

There was also a concern from the beginning of language documentation for supporting speakers and communities who wish to retrieve, revitalise or maintain their languages by providing documentation corpora that could be connected to revitalization work (but see Section 5 below). Also playing a role were advances in information, media, communication and archiving technologies (see Nathan 2010a, 2010b and Section 4) which made possible the collection, analysis, preservation and dissemination of documentary corpora in ways which were not feasible previously. Language documentation also paid attention to the rights and needs of language speakers and community members, and encouraged collaborative

---

2.   Himmelmann (2008: 339) argues that the 'trigger … was a short presentation by Johannes Bechert … at the fourteenth International Congress of Linguists in East Berlin in 1987. … [and] a motion drafted by Christian Lehmann, which was presented to the business meeting of the Comité International Permanent des Linguistes (CIPL) … [urging] the committee to take action with the goal of bringing the issue of language endangerment to the attention of professional linguists and the general public'. Also important, especially in North America, was Hale et al. (1992).

approaches that would include their direct involvement in the documentation and support of their own languages (see Grinevald 2003; Austin 2010; Yamada 2007).

A concurrent and supporting development was the availability of extensive new funding resources for research on endangered languages from several sources, and the requirements of these funders to adopt a documentary perspective and to archive the recorded data and analyses. The new funders included the Endangered Languages Documentation Programme (ELDP)[3] at SOAS (established in 2002 by Arcadia Fund, it has now provided around 350 documentation grants), the Volkswagen Foundation DoBeS[4] project (which ran from 2001 to 2014 and funded 80 projects), and the Documenting Endangered Languages (DEL)[5] inter-agency programme of the National Science Foundation and the National Endowment for the Humanities (established 2005, it has funded 320 projects to date). Other smaller sources also emerged (the Endangered Language Fund (ELF),[6] Foundation for Endangered Languages (FEL),[7] Gesellschaft für bedrohte Sprachen (GBS)[8] and Unesco[9]) and have made more modest grants supporting scores of projects, many of which are community-based. This new funding influenced the topics that linguists (and others) chose to research, and the research methods they employed (see Sections 2 and 4 below).

The broader impact on the field of linguistics can be seen in the development of:

– academic journals specialising in language documentation topics (*Language Documentation and Conservation*,[10] *Language Documentation and Description*[11]), and special issues of other linguistics journals dedicated to documentation and revitalisation (e.g. Volume 34/4 (2013) of the *Journal of Multilingual and Multicultural Development*[12]);

---

3.   <http://www.hrelp.org/grants> (11 March 2015).

4.   <http://dobes.mpi.nl/dobesprogramme> (12 March 2015).

5.   <http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=12816> (27 March 2015).

6.   <http://www.endangeredlanguagefund.org/> (27 March 2015).

7.   <http://www.ogmios.org/index.php> (27 March 2015).

8.   <http://www.uni-koeln.de/gbs/> (27 March 2015).

9.   <http://www.unesco.org/new/en/culture/themes/endangered-languages> (27 March 2015).

10. <http://nflrc.hawaii.edu/ldc/> (14 March 2015).

11. <http://www.elpublishing.org> (10 March 2015).

12. <http://www.tandfonline.com/toc/rmmm20/34/4#.VRTkmvmUeSo> (27 March 2015).

- specialist conferences, such as the International Conference on Language Documentation and Conservation held biennially in Hawaii[13] and the Language Documentation and Linguistic Theory (LDLT)[14] conference held biennially since 2007 at SOAS;
- workshops and training courses, including the Summer Institutes of CoLang/ InField[15] run biennially in the United States since 2008, summer schools of the 3L consortium (Leiden-London-Lyon) that also commenced in 2008, and the DocLing[16] training course held annually at Tokyo University of Foreign Studies since 2008;
- specialist MA and PhD programmes at SOAS[17] (Austin 2008), University of Hawaii,[18] and the increasing introduction of documentation topics in undergraduate and postgraduate Linguistics programmes elsewhere;
- a growing number of book publications on topics related to language documentation (for an annotated bibliography see Austin 2013);
- increased attention among linguists with a range of interests, objectives and theoretical persuasions to issues of data quality, portability, data citation, glossing standardization, and data sources (including elicitation, translation, story boarding, naturalistic observations, and experimentation).

Himmelmann (2006: 15) identified five major characteristics of language documentation that he proposed would distinguish it from other approaches to the study of human languages:

- *focus on primary data* – language documentation concerns the collection and analysis of an array of primary language data to be made available for a wide range of users (further elaborated in Himmelmann 2012);
- *explicit concern for accountability* – access to primary data and representations of it makes evaluation of linguistic analyses possible and expected;
- *concern for long-term storage and preservation of primary data* – language documentation includes a focus on archiving in order to ensure that documentary materials are made available to potential users now and into the distant future;

---

13. <http://icldc-hawaii.org/> (27 March 2015).

14. <http://www.hrelp.org/events/> (27 March 2015).

15. <http://www.alaska.edu/colang2016/charter/> (27 March 2015).

16. <http://www.aa.tufs.ac.jp/en/training/fieldling-ws/docling> (27 March 2015).

17. <https://www.soas.ac.uk/linguistics/programmes/malangdocdesc/> (27 March 2015).

18. <http://ling.hawaii.edu/> (27 March 2015).

- *work in interdisciplinary teams* – documentation requires input and expertise from a range of disciplines and is not restricted to linguistics alone;
- *close cooperation with and direct involvement of the speech community* – language documentation requires active and collaborative work with community members both as producers of language materials and as co-researchers.[19]

The application of these principles results, according to Himmelmann (1998, 2002, 2006), in the creation of a record of the linguistic practices and traditions of a speech community together with information about speakers' metalinguistic knowledge of those practices and traditions. This is achieved by systematic recording, transcription, translation and analysis of a variety of spoken (and written) language samples collected within their appropriate social and cultural context. Analysis within language documentation under this view is aimed at making the records accessible to a broad range of potential users which includes not only linguists but also researchers in other disciplines, community members and others, who may not have first-hand knowledge of the documented language. The record is also intended for posterity (and hence should be preservable and portable, in the sense of Bird & Simons 2003), and so some level of processing is required. There is a need for systematic recording of metadata (data about the data) to make the archived materials understandable, findable, preservable and usable.

The core of a language documentation defined in this way was generally understood to be a corpus of audio and/or video materials with time-aligned transcription, annotation, and translation into a language of wider communication (Schulze-Berndt 2006), and relevant metadata on context and use of the materials. Woodbury (2003) argued that the corpus will ideally cover a diverse range of genres and contexts, and be large, expandable, opportunistic, portable, transparent, ethical and preservable. Austin (2006a, 2008, 2010) proposes that there are five activities (not necessarily sequential) which are identifiable in this documentation approach and which contribute to corpus creation, analysis, preservation and dissemination:

- *recording* – of media and text (including metadata) in context;
- *transfer* – to a data management environment;
- *adding value* – the transcription, translation, annotation and notation and linking of metadata to the recordings;

---

19. Issues concerning communities, collaboration and ethics of research have been an ongoing thread in papers published in the journal *Language Documentation and Conservation* over a number of years.

- *archiving* – creating archival objects and assigning them access and usage rights;
- *mobilisation* – creation, publication and distribution of outputs, in a range of formats for a range of different users and uses.

2.   Best practices, tools and models

The establishment of the DoBeS project in 2001 saw the emergence of a unified 'DoBeS model' for language documentation that the funded projects were expected to adopt.[20] This included specifications for archival storage, recommendations about recording and analysis formats, and the development of new software tools to assist with audio and video annotation (such as ELAN[21]), and the creation and management of metadata (various IMDI tools[22]). Researchers affiliated with DoBeS also proposed general principles (or 'best practice') for language documentation, such as sampling (to meet Himmelmann's desideratum that the documentary record should be 'representative', see Seifart 2008), data collection methods (Lüpke 2009) and a typology of data types (Himmelmann 2012).

Definition of best practice, standards, tools and models was also a central goal of the E-MELD project[23] funded by the National Science Foundation which ran from 2001 to 2006 aiming to develop recommendations for metadata, annotation markup, language identification and linguistic ontology (essentially the sets of labels employed in interlinear glossing). This resulted in a series of papers[24] defining formats for lexical entries (Bell & Bird 2000), interlinear text (Bird & Liberman 2001; Bowe et al. 2003), paradigms (Penton et al. 2004) and a generalised ontology for glossing (Farrar et al. 2002; Farrar & Langendoen 2003a, b). E-MELD set up a 'School of Best Practices' (Aristar 2003; Aristar-Dry 2004)[25] with case studies, a reference list of readings and tools, and a classroom 'designed to offer "lessons" and tutorials which explain the recommendations of best practices'.

---

20. <http://dobes.mpi.nl/dobesprogramme and http://www.mpi.nl/corpus/a4guides/a4-guide-dobes-format-encoding.pdf> (10 March 2015).

21. <https://tla.mpi.nl/tools/tla-tools/elan/> (10 March 2015).

22. <http://www.mpi.nl/IMDI/> (10 March 2015).

23. <http://emeld.org/> (10 March 2015).

24. <http://emeld.org/documents/index.cfm#loc-papers> (10 March 2015).

25. <http://emeld.org/school/index.html> (10 March 2015).

Probably the most ambitious attempt to define best practice and what would constitute a complete documentation of a language is to be found in CELP 2007, which attempted to define everything that an adequate documentation should cover: all the basic phonology, morphology, syntactic constructions (in context), and provide a lexicon covering all the basic vocabulary and important areas of special expertise in the culture, with at least glosses for all words/morphemes in the corpus, plus a full range of textual genres and registers. It offered a set of 'accounting standards' to determine adequacy, including quantitative measures such as a figure of 10,000 items for a lexicon, and a text corpus of one million words (around 1200 hours of recorded speech). Other qualitative measures were suggested such as the notion that research on an endangered language is completed 'when nothing new is coming up in non-elicited material and when any apparent lacunae in the phonological system can be shown to be real and not an accident of data collection'.

It is doubtful if linguists would ever suggest it is possible to qualitatively and quantitatively determine when a research project is 'complete' for non-endangered languages, yet this is precisely what was suggested for language documentation.

Both DoBeS and E-MELD were influential in encouraging linguists to begin to pay attention to data types, data structures, analytical processes and workflows, together with preservability and transparency, however the notion that there was a 'documentation model' or a 'best practice' (or a small number of 'best practices') was questioned by some researchers, beginning around 2004.

3.   Critical responses

The role of archives in defining the goals and values of language documentation was challenged by Nathan (2004) who introduced the term 'archivism' to describe the idea that quantifiable properties such as recording hours, data volume, file parameters, and technical desiderata like 'archival quality' and 'portability' could be reference points in assessing the aims and outcomes of language documentation. He argued that these should not be measures of quality of a documentation project, and that there had been a lack of discussion of research methodology among language documenters, including about what such quality measures might be.

Nathan and Austin (2004) addressed the issue of metadata and argued that all value-adding that researchers provide for the audio or video records they make should be understood as metadata, and that it should be as rich as possible and designed for the documentation purpose at hand. This means that metadata should not be constrained by specifications in the form of an 'ontology' or standard

minimal set (such as that proposed by OLAC[26]). The need for richer metadata and meta-documentation (documentation of the language documentation) was further elaborated on by Austin (2009, 2013) – see also Gawne et al. (2015).

Two important issues for the definition of language documentation were raised in 2006, namely the difference between documentation and description, which was considered fundamental in Himmelmann's seminal paper (see quotation in 1 above), and the approach to audio recording within documentation. Austin (2006b), revised and published as Austin and Grenoble (2007), noted that, as Himmelmann (1998) made clear, language documentation and description differ in terms of their goals, areas of interest, research methods, workflows, and outcomes. Language description focusses on languages as sets of structures and systems, and typically aims to produce grammars, dictionaries, and collections of texts, the intended audience of which is usually linguistics specialists. By contrast, documentation is discourse-centered: its primary goal is the representation of a range of instances and types of language use in their social and cultural context. Although description may draw on a corpus, it involves analysis of a different order, oriented to providing an understanding of language at a more abstract level, as a system of elements, rules, and constructions. Austin and Grenoble (2007:22) challenged this sharp separation of description and documentation and argued that:

> [d]ocumentation projects must rely on the application of theoretical and descriptive linguistic techniques in order to ensure that they are usable (i.e. have accessible entry points via transcription, translation and annotation), as well as to ensure that they are comprehensive. It is only through linguistic analysis that we can discover that some crucial speech genre, lexical form, grammatical paradigm or sentence construction is missing or under-represented in the documentary record. Without good analysis, recorded audio and video materials do not serve as data for any community of potential users.

In terms of workflow, they also differ. For description, linguistic knowledge and decision-making is applied to some event in the real world to make an inscription (e.g. an audio recording) that is not itself of interest but serves as a source which can then be selected, analysed and systematised in order to create analytical representations, typically in the form of lists, summaries and analyses (e.g. statements about phonology, morphology or syntax). It is these representations which are the main focus of interest and which are then presented and distributed to users, typically other linguists. For documentation, linguistic and cultural knowledge and documentary techniques are applied to some event in the real world to make

---

26. <http://www.language-archives.org/OLAC/olacms.html> (10 March 2015).

an inscription (audio or video recording) that recapitulates aspects of the original event (such as social or spatial relationships – see Nathan 2010a) and is itself a focus of interest (e.g. for archiving, preservation and distribution). The documentary researcher adds value to the inscription by making decisions and applying linguistic and other knowledge to create representations, typically in the form of transcriptions, translations and annotations. These representations are the second major focus of interest and will be archived and/or mobilized and distributed. The same representations could, of course, also be the input to the selection and analytical procedures of description, thereby linking the descriptive outcomes to the documentary corpus. From this viewpoint, documentation and description are complementary activities with complementary goals, methods and outcomes.

Nathan (2006) argued that despite the expressed concern by language documenters for recording language in its social and cultural context, many researchers took an unscientific approach to audio recording in particular, ignoring issues such as spatiality and microphone selection and placement. He extended this critique in Nathan (2009, 2010a) and argued for the need to establish an epistemology for audio recording within language documentation.

A broader critique of documentation and contemporary endangered languages research can be found in Dobrin et al. (2007) who identify and highlight tendencies towards objectification of languages, and reliance on familiar qualitative metrics to measure quality, progress and value. More specifically, they argue that 'subtle and pervasive kinds of commoditisation (reduction of languages to common exchange values) abound, particularly in competitive and programmatic contexts such as grant-seeking and standard-setting where languages are necessarily compared and ranked'. Bowern (2011: 468) also points to commoditisation and suggests that 'community members report sometimes feeling that the linguist comes in, reifies the language, turns it into a commodity, and then takes it away.'

Dobrin et al. (2007) echo Nathan (2004) in pointing to archivism as problematic, and join Nathan (2006) in arguing that documentary linguists show little or no knowledge about recording arts, including microphone types, properties and placement, even though microphone choice and handling is the single greatest determiner of audio recording quality. They also note that evidence from archival deposits shows that video tends to be poorly used by documentary linguists, with video recordings being made without reference to articulated hypotheses, goals, or methodology, simply because the technology is available, portable and relatively inexpensive. Finally, in contrast to earlier approaches, they point to diversity as an important aspect of language documentation. As researchers respond to the unique and particular social, cultural and linguistic contexts within which the languages they are studying are spoken or signed, actual documentation projects, as evidenced by grant project proposals and materials deposited in archives,

show a diversity of approaches, techniques, methodologies, skills and responses. In the last 10 years we also find an increasing diversity of materials that can be included in corpora, so that alongside the traditional field interviews, observations, experiments and narrative collections that have been the bread and butter of documentation and description, we also find materials, much of them created by native speakers, from YouTube uploads, Twitter feeds, Facebook posts, blogs, email, chat, Skype calls, and local pedagogy developed for revitalization. Similarly, the outcomes of documentation are increasingly diverse so that alongside books, papers and archive deposits, today research projects are also generating YouTube uploads,[27] Twitter and Facebook posts, blogs,[28] multimedia (such as *Gayarragi Winangali*,[29] and mobile apps (such as *Ma! Iwaidja*[30]). Rather than aiming for comprehensiveness or representativeness, research funded recently by ELDP for example, rather shows specificity, focussing on topics such as traditional song in its diaspora context, language use by blacksmiths, bark cloth making, libation rituals, fishing practices, child language, interactive speech, and ethnobotany (projects funded in 2012 and 2014).[31]

In a recent handbook, Woodbury (2011:159) presents a definition of language documentation which reflects this shift away from representative samples towards more specific goals as 'the creation, annotation, preservation and dissemination of transparent records of a language'. He also identifies some gaps in the earlier conceptions of documentation, especially because 'language encompasses conscious and unconscious knowledge, ideation and cognitive ability, as well as overt social behaviour' (ibid.). The role of ideologies of language structure and use, attitudes of speakers to their and others' speech, and the relationships of beliefs and attitudes to actual performance in the world are only beginning to be addressed by documentary linguists (see Austin & Sallabank 2014). As Woodbury (2011:160) notes, 'humans experience their own and other people's languages

---

27. For example, Anthony Jukes' subtitled video on Minahasan food and cooking methods at <https://www.youtube.com/watch?v=wVy2QsFqdYI> (9 June 2015); see also <https://www.youtube.com/watch?v=hqNQ-z9sIBw> for further details.

28. For example, Austin's Dieri blog at <http://www.dieriyawarra.wordpress.com> (15 March 2015).

29. <http://www.dnathan.com/projects/gw/> (27 March 2015).

30. <http://www.iwaidja.org> (15 March 2015).

31. <http://www.hrelp.org/grants/projects/index.php?year=2012>   <http://www.hrelp.org/grants/projects/index.php?year=2014> (14 March 2015). Note that there has not been a complete shift away from the 'whole language documentation' approach with quite a number of funded projects still taking such an approach.

viscerally and have differing stakes, purposes, goals and aspirations for language records and language documentation'.

Woodbury (2011) has also highlighted a need to develop a theory of documentary corpora (covering the principles by which a particular corpus 'hangs together'), as well as a need for accounts of individual documentation project designs. Austin (2013) extends this to a general call for reflexive meta-documentation of their work by researchers concerning their documentary models, processes and practices. This would include: the identity of stakeholders and their roles; the attitudes and ideologies of language consultants and the communities within which they are located (towards their languages as well as the documenter and documentation project[32]); the relationships between researchers, research project participants and the wider community; the goals and methodology adopted by the project, including research methods and tools (see Lüpke 2009); corpus theorization (Woodbury 2011); theoretical assumptions embedded in annotation and translation (e.g. in abbreviations, glosses); and considerations of the potential for a project to contribute to revitalization. In addition, it is important to know the biography of the project, including background knowledge and experience of the researcher and main consultants (e.g. how much fieldwork the researcher had done at the beginning of the project and under what conditions, what training the researcher and consultants had received). Austin (2013) suggests that such meta-documentation can draw upon knowledge and practices in other disciplines (such as social and cultural anthropology, archaeology, archiving and museum studies), and from considerations that surface in the interpretation of past documentations (of legacy materials). The many parallels between language documentation and ethnomusicology in terms of these and other topics are explored in detail by Grant (2014).

Austin and Sallabank (2015) point out that the early emphasis on 'compiling a representative and lasting multipurpose record of a language' has led documenters to focus on defining and describing individual languages in isolation with a narrow attention to what Woodbury (2011: 177) calls 'the ancestral code', rather than documenting dynamic language practices and real-life interactions in their sociolinguistic context (see also Sugita 2007; Amery 2009; Childs et al. 2014). By definition, endangered languages do not exist in isolation but are always spoken in relationships with other languages, varieties, codes, styles, registers, etc., in a complex linguistic ecology (Haugen 1972; Mühlhäusler 1992, 2000; Calvet 2006). Grenoble (2011) has argued that linguists should aim to document language ecologies, not just what they define as individual languages or varieties (the ancestral

---

32. See Kroskrity (2015) for an example relating to a dictionary project.

code approach). At the very least they should pay attention to multilingual repertoires, mixed codes, the sociolinguistic and structural effects of contact, and language variation and change (Lüpke & Storch 2013). Gullberg (2012) has explored the interplay between multilingualism and multimodality, arguing that 'language documentation data has the potential to inform theoretical and empirical studies of linguistics, bilingualism and multimodality in entirely new ways, and, conversely, that documentation work would benefit from taking the bilingual and multimodal nature of its data into account' (Gulberg 2012: 46).

It is also important to consider extra-linguistic factors such as language attitudes and ideologies (Sallabank 2013; Austin & Sallabank 2014). The dominant model of language documentation from 1995 to 2010 could be described as 'saving the morphemes two-by-two' in a 'Noah's arc(hive)', salvage-linguistics approach which reflects a purist notion of single languages in isolation. From 2010, for at least some language documenters, the approach has become more particular, dynamic, pluralistic and socially-engaged.

4.   Developments in archiving

The rise of language documentation has also seen the development of a number of internet-accessible digital archives focusing in particular on the preservation of materials on endangered languages. These include DoBes in the Netherlands,[33] Paradisec in Australia,[34] Pangloss in Paris,[35] the California Archive in Berkeley,[36] AILLA in Texas,[37] and ANLA in Alaska.[38]

One of the most dramatic developments of the 21st century has been the rise of social network models on the internet (so-called Web 2.0) that aim to link people rather than documents, with a focus on interaction and collaboration instead of passive downloading and viewing of content. These new models have been taken up in the last 10 years by some language documentation archives (such as ELAR[39] at SOAS) leading to what Nathan (2010b) calls 'Archives 2.0'.

---

33. <http://dobes.mpi.nl/> (9 June 2015).

34. <http://paradisec.org.au/> (9 June 2015).

35. <http://lacito.vjf.cnrs.fr/pangloss/presentation_en.htm> (9 June 2015).

36. <http://cla.berkeley.edu/> (9 June 2015).

37. <http://www.ailla.utexas.org/site/welcome.html> (9 June 2015).

38. <https://www.uaf.edu/anla/> (9 June 2015).

39. <http://www.elar-archive.org> (10 March 2015).

Traditionally, archiving has focused heavily on preservation (and on cataloguing and standards – see Section 3 above), however language documentation raises a number of new methodological challenges, especially in relation to endangered languages where speakers 'tend to use their language more and more to speak of private, local, sensitive and secret matters. So the primary data of documentary linguistics maximises the likelihood of including content that can cause discomfort or harm to the recorded speakers' (Nathan 2014: 191) or their families and descendants. Thus documentation corpora often contain ritual or sacred material that may be restricted in terms of who can be exposed to them, as well as gossip which may contain references to private knowledge or events. As a result, language documentation archives need powerful but flexible access management that is transparent, easy to understand, and able to be changed as circumstances develop. The basis for access will be via relationships between the providers of the materials (archive depositors and the stakeholders they work with) and those who wish to use them. Beginning in 2005, the ELAR archive at SOAS developed a richly articulated system of 'access protocols' designed to formulate and implement speakers' rights and sensitivities, together with rigorous methods and processes for controlled access to the archival materials. Each resource is assigned one of five levels of access: U (open to all registered users), R (for registered researchers only), C (for community members only) and S (for subscribers who negotiate access with the depositor), X (closed to all but the depositor). Registered users are then categorized by archive staff and their access to particular materials depends on their status (e.g. they are R by virtue of being associated with an academic programme, and/or C because they explain that they have links to a particular community[40]) and the access type of the materials they wish to use. A similar access protocol system is in use by TLA, *The Language Archive*, at the Max Planck Institute for Psycholinguistics (which includes the DoBeS endangered languages archive).[41] Endangered language archiving thus requires a special response to the well-publicised movement for complete open access that is current in much other academic research and publication.

In this view, such an archive can also be seen as a place for establishing and transacting relationships and sharing, and Web 2.0 models provide a technology for instantiating this. The general model of the ELAR archive is presented by Nathan (2010b) as in Figure 1.

There are several other archiving developments that have been pioneered by ELAR in the last 10 years. The first, called *progressive archiving*, sees archiving

---

40. This can be one of the most difficult and complex statuses for an archive to determine.

41. <http://dobes.mpi.nl/access_registration/> (10 March 2015).
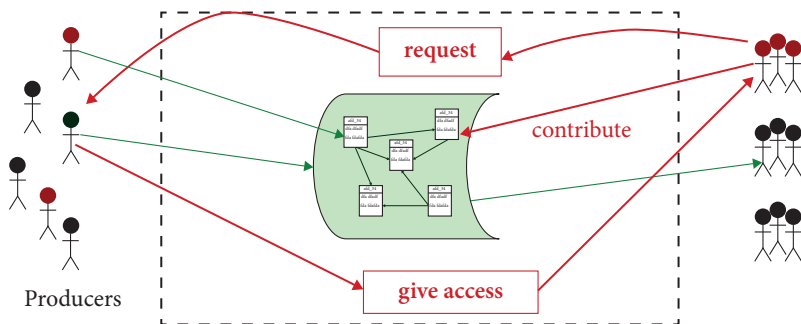
Figure 1. ELAR Archive 2.0 model

as a whole-of-project relationship: depositor accounts are established at the beginning of a research project, and researchers add and manage or update their materials over time, as well as managing and engaging in interactions with the curators and users. Secondly, ELAR have developed a web accessible archive interface that has been designed to provide contextualization, different degrees of presentation for different projects, and ease of navigation for users. The interface directly reflects the interests and needs of the materials providers and the users, rather than being, for example, a unified tree structure across the whole collection, as other archives such as DoBeS and AILLA. Thirdly, ELAR has promoted increased participation so that users can negotiate access to particular materials and bookmark their favourites, while depositors can negotiate access requests and monitor usage. A communication channel has also been established in order for both groups to exchange and share information. Nathan (2014) gives examples of these exchanges and how they can lead to creative outcomes and collaborations between researchers and members of the community of users.

Possible future developments in endangered languages archiving may include community curation of archived materials (Linn 2014), participant identification and expression of rights (Garrett 2014), and the creation of new kinds of outputs that draw upon a range of materials drawn from several collections within the archive (just as museums and galleries choose, select and exhibit their resources for educational or other purposes – see Holton 2014). The overall flavour of archiving in the last five years has changed from finality and completeness to being open and evolutionary. These developments also raise questions for archives about what a 'deposit' or 'depositor' really is, and recast archives as providers of services within a revised, holistic concept of language documentation.

5.   Language documentation and revitalization

The term 'language revitalization' is used to describe principles and activities aimed at increasing the number of users of a language, and/or the range of domains within which it is used (Fishman 1991, 2001; Hinton & Hale 2001; Hornberger 2010; Hinton 2011; Romaine 2007; Grenoble & Whaley 2006). It has been in operation for more than 20 years longer than language documentation as its origins go back to community-based activities by Māori in New Zealand in the 1970s (Spolsky 1989, 2003; Bentahila & Davies 1993) and by other groups such as North American indigenous people (Niedzielski 1992; Kapono 1995; Hinton 1993, 2002, 2013), and European minorities such as the Catalan, Welsh and Basque.

The relationship between language documentation and language revitalization is a rather complex one, and is explored in some detail in Austin and Sallabank (2015). For many language documenters revitalization has been seen as a waste of resources, a viewpoint connected to the 'language-as-system' ideology that sees linguistic data as the only thing worth collecting and preserving, in contrast to 'linguistic social work' (Newman 2003: 6; see also Dimmendaal 2004: 84 and Blench 2008: 153).[42]

Although documentation defined itself from the beginning as a field that set out to create a multipurpose record for a wide range of users, including community members, language revitalization has been treated as a simple 'technical add-on' that involves creation of orthographies, dictionaries, sub-titled videos, and primers and multimedia, including websites, rather than as a field of research or activity that requires theoretical and applied knowledge. This view was also strongly supported by the funding agencies (including ELDP, Volkswagen Stiftung, NSF-NEH), who excluded revitalization-oriented projects from grants and severely limited the amount of money that could be included for revitalisation materials creation or 'community publication' of research results.

Much of the material that has ended up in language documentation archives is unsuitable for revitalization for a variety of reasons, including inappropriate genres or topics, recordings and analyses in difficult to access archival formats that require specialised software (such as ELAN or FLEx), or glossed and translated into languages such as English that have little or no place in the local linguistic ecology. Documentation is also heavily biased towards the performances of older fluent speakers, resulting in language that may be too fast, heavily context dependent and include slurring or elisions, or even be affected by physiological factors (not least of which may be lack of teeth). Few, if any, documentary

---

42. Newman's views were repeated and further elaborated in Newman (2013); see also the response by Whalen (2013).

corpora include samples of children's ordinary language use[43] or learner-directed speech; in addition, as noted by Cope (2014), documentary linguists are not trained in pedagogical materials design, and applied linguists are rarely included in language documentation teams. The relationship between language documentation and revitalization has thus varied from avoidance or subordination to, at best, only an indirect connection (Sallabank 2012). There is a need for much more exploration and development of this area in the future (see also Austin & Sallabank 2015 for further discussion).

6. Documentation and academia

The development of language documentation as a field with its own principles and practices appeared to many researchers in its foundation period at the end of the 20th century to offer an opportunity to change the socio-political academic balance between fieldworkers and so-called 'armchair linguists' (typologists, theoreticians) (Fillmore 1992; Aikhenvald 2007: 4; Crowley 2007: 11–13) by providing a foundation (theory, best practices) for corpus creation, data collection and analysis. Many perceived that fieldwork and language description were in a subordinate sociological position. Newman (2009: 124)[44] states explicitly that 'theoreticians belittle descriptivists as linguistically second-class citizens'[45]), and hoped that language documentation and the work of corpus creation and associated activities would raise their status in academic linguistics. Indeed, lobbying by documenters and others led in 2010 to the Linguistic Society of America 'Resolution Recognizing the Scholarly Merit of Language Documentation' which states that:

> [a] shift in practice has broadened the range of scholarly work to include not only grammars, dictionaries, and text collections, but also archives of primary data, electronic databases, corpora, critical editions of legacy materials, pedagogical works designed for the use of speech communities, software, websites, or other

---

43. An exception is the DoBeS Chintang/Puma project – see http://dobes.mpi.nl/projects/chintang/, accessed 9 June 2015. There is incidental children's language material in the ELAR archive, such as children's retellings of the Frog Story book (Mayer 1969), however this material has not been systematically collected.

44. Originally published in 1992.

45. Newman (2009: 124) considers this to be an 'unintended consequence of Chomsky's (1964) hierarchy of levels of adequacy in grammar, namely, from the bottom up, observational adequacy – "A grammar that aims for observational adequacy is concerned merely to give an account of the primary data" (p. 63, italics mine) –, descriptive adequacy, and explanatory adequacy.'

digital media. The products of language documentation and work supporting linguistic vitality are of significant importance to the preservation of linguistic diversity, are fundamental and permanent contributions to the foundation of linguistics, and are intellectual achievements which require sophisticated analytical skills, deep theoretical knowledge, and broad linguistic expertise.

The resolution 'support[ed] the recognition of these materials as scholarly contributions to be given weight in the awarding of advanced degrees and in decisions on hiring, tenure, and promotion of faculty'. In addition, the resolution encouraged 'the development of appropriate means of review of such works so that their functionality, import, and scope can be assessed relative to other language resources and to more traditional publications'.

To date, criteria for this kind of review of documentary corpora, or examples of such reviews (parallel, say, to book reviews), have not appeared. In the five years since this resolution was passed there still remains what we can call an 'output gap': traditional products of language description and typological and theoretical research (grammars, book chapters, journal articles) are understood and accorded value in determining promotion, award of tenure and in decision making about new job appointments, but the newer outputs in the form of digital archival deposits, multimedia products, and pedagogical materials for revitalization are either not valued or discounted.

According to Thieberger (2012) similar discussions have taken place in Australia beginning in 2011 between the Australian Linguistic Society (ALS) and the Australian Research Council (ARC), and 'although the ARC accepted that curated corpora could legitimately be seen as research output, it would be the responsibility of the ALS (or the scholarly community more generally) to establish conventions to accord scholarly credibility to such products'. He reports on proposals for a possible review procedure but recognizes that 'the question of what criteria to use in evaluating a corpus is more problematic'. For some suggestions for criteria see Thieberger (2012) and Thieberger et al. (2012); again no action appears to have been taken to date to actually implement these Australian proposals.

In my view, to address this output gap, there is a need for experimentation and the development of new genres, so far unfamiliar to linguists, that link and contextualise analytical outputs and the archival corpus. These could include ethnographies of documentation project designs, accounts of data collection (cf. the genre of research publications in archaeology called 'field reports'), finding-aids to corpus collections, or 'exhibitions' or 'guided tours' of archival deposits (along the lines of exhibitions and associated products regularly mounted by museums to display parts of their collections, see also Woodbury 2014). Similarly, reviews of corpora or these new kinds of writing could also be attempted.

There has been a very recent development in Linguistics of free online open access publication platforms (e.g. *Language Science Press*, established in April 2013,[46] and *EL Publishing*, launched in July 2014),[47] with all the usual academic requirements such as double-blind reviewing and professional editing, design and layout. While *Language Science Press* publishes digital versions of traditional books, *EL Publishing* has set out to provide and encourage new opportunities for language documenters to publish multimedia and the other innovative types of output mentioned above. It remains to be seen whether these opportunities will be taken up by practitioners, and whether they will go some way to addressing the output gap in the future.

7.   Conclusions

The past 20 years has seen the emergence and gradual development of a new field of research called 'documentary linguistics' or 'language documentation' which has concentrated on recording, analysing, preserving and disseminating records of languages in use in ways that can serve a wide range of constituencies, particularly the language communities themselves. In the early period of its development there was a concentration on defining a model for language documentation and specifying best practices, tools and analytical categories, however the past 10 years have seen a shift in perspective responding to criticism of these early concerns. Today, there is more recognition of diversity of contexts, goals, methods and outcomes of language documentation, and indications of the introduction of social models of research, especially in the area of archiving. Much work remains to be done however, to engage better with language revitalization and to establish reliable and replicable measures for evaluating the quality, significance and value of language documentation research so that its position alongside such sub-fields as descriptive linguistics and theoretical linguistics can be assured and enhanced.

References

Aikhenvald, A.Y. 2007. Linguistic fieldwork: Setting the scene. *Sprachtypologie und Universalienforschung* 60(1): 3–11.

Amery, R. 2009. Phoenix or relic? Documentation of languages with revitalization in mind. *Language Documentation and Conservation* 3(2): 138–148.

---

46. <http://langsci-press.org/index> (9 June 2015).

47. <http://www.elpublishing.org> (1 March 2015).

Aristar, A.R. 2003. *The school of best practice*. London: SOAS Workshop on Archives for Endangered Languages, November 21–22.

Aristar-Dry, H. 2004. E-MELD School of best practices in digital language documentation. Detroit: Presentation at *E-MELD Conference 2004: Workshop on linguistic databases and best practice*, July 15–18.

Austin, P.K. 2006a. Data and language documentation. In *Essentials of Language Documentation* [Trends in Linguistics. Studies and Monographs 178], J. Gippert, N. Himmelmann & U. Mosel (eds), 87–112. Berlin: Mouton de Gruyter.

Austin, P.K. 2006b. *Defining language documentation*. Washington DC: Paper presented at the Georgetown University Roundtable on Linguistics, March.

Austin, P.K. 2007. Survival of languages. In *Survival: Darwin College Lectures*, E.F. Shuckburgh (ed.), 80–98. Cambridge: CUP.

Austin, P.K. 2008. Training for language documentation: Experiences at the School of Oriental and African Studies. In *Documenting and Revitalising Austronesian Languages* [Language Documentation and Conservation Special Publication 1], M. Florey & V. Rau (eds), 25–41. Hawaii HI: University of Hawaii Press.

Austin, P.K. 2009. Meta-documentary linguistics. Kioloa: Paper given at the *Aboriginal languages workshop*, March 2009.

Austin, P.K. 2010. Current issues in language documentation. *Language Documentation and Description* 7, 12–33. London: SOAS.

Austin, P.K. 2013. Language documentation and meta-documentation. In *Keeping Languages Alive: Documentation, Pedagogy and Revitalization*, S. Ogilvie & M. Jones (eds), 3–15. Cambridge: CUP. doi:10.1017/CBO9781139245890.003

Austin, P.K. 2014. Language documentation in the 21st century. *JournaLIPP* 3: 57–71.

Austin, P.K. & Grenoble, L. 2007. Current trends in language documentation. *Language Documentation and Description* 4, 12–25. London: SOAS.

Austin, P.K. & Sallabank, J. 2014. *Endangered Languages: Ideologies and Beliefs in Language Documentation and Revitalization*. London: British Academy. doi:10.5871/bacad/9780197265765.001.0001

Austin, P.K. & Sallabank, J. 2015. Language documentation and language revitalization: Partners or just good friends? Ms, SOAS.

Bell, J. & Bird, S. 2000. Preliminary study of the structure of lexicon entries. Philadelphia: Proceedings of the *Workshop on Web-Based Language Documentation and Description*.

Bentahila, A. & Davies, E.E. 1993. Language revival: Restoration or transformation? *Journal of Multilingual and Multicultural Development* 14(5): 355–374. doi:10.1080/01434632.1993.9994542

Bird, S. & Liberman, M. 2001. A formal framework for linguistic annotation. *Speech Communication* 33(1–2): 23–60. doi:10.1016/S0167-6393(00)00068-6

Bird, S. & Simons, G. 2003. Seven dimensions of portability for language documentation and description. *Language* 79(3): 557–582. doi:10.1353/lan.2003.0149

Blench, R. 2008. Endangered languages in West Africa. In *Language Diversity Endangered*, M. Brenzinger (ed.), 140–162. Berlin: Mouton de Gruyter.

Bowe, C., Hughes, B. & Bird, S. 2003. Towards a general model for interlinear text. Proceedings of *EMELD-03*, 12 August, 2013. <http://emeld.org/workshop/2003/bowbadenbird-paper.html>

Bowern, C. 2011. Planning a language documentation project. In *The Cambridge Handbook of Endangered Languages*, P.K. Austin & J. Sallabank (eds), 459–482. Cambridge: CUP. doi: 10.1017/CBO9780511975981.023

Calvet, J.-L. 2006. *Towards an Ecology of World Languages*. Cambridge: Polity.

CELP. 2007. Adequacy of documentation. Document circulated at the *January 2007 meeting of the Linguistic Society of America Committee on Endangered Languages and Their Preservation*.

Childs, T., Good, J. & Mitchell, A. 2014. Beyond the ancestral code: Towards a model for sociolinguistic language documentation. *Language Documentation and Conservation* 8: 168–191.

Chomsky, N. 1964. Current issues in linguistic theory. In *The Structure of Language*, J.A. Fodor & J.J. Katz (eds), 50–118. Englewood Cliffs NJ: Prentice Hall.

Coelho, G. 2005. Language documentation and ecology: Areas of interaction. In *Language Documentation and Description* 3, P.K. Austin (ed.), 63–74. London: SOAS.

Cope, L. (ed.). 2014. *Applied Linguists Needed: Cross-disciplinary Networking in Endangered Language Contexts*. Abingdon: Routledge.

Crowley, T. 2007. *Field Linguistics: A Beginners Guide*. Oxford: OUP.

Crystal, D. 2000. *Language Death*. Cambridge: CUP. doi: 10.1017/CBO9781139106856

Dimmendaal, G. 2004. Capacity building in an African context. In *Language Documentation and Description* 2, P.K. Austin (ed.), 71–89. London: SOAS.

Dobrin, L., Austin, P.K. & Nathan, D. 2007. Dying to be counted: Commodification of endangered languages in documentary linguistics. *Language Documentation and Description* 6, 37–52. London: SOAS.

Eisenbeiss, S. 2005. Psycholinguistic contributions to language documentation. *Language Documentation and Description* 3, 106–140. London: SOAS.

Farrar, S. & Langendoen, D.T. 2003a. A linguistic ontology for the semantic web. *GLOT International* 7(3): 97–100.

Farrar, S. & Langendoen, D.T. 2003b. Markup and the GOLD ontology. *Proceedings of EMELD-03*, 12 August 2013. <http://staff.washington.edu/farrar/documents/inproceedings/FarLang03a.pdf>

Farrar, S., Lewis, W.D. & Langendoen, D.T. 2002. A common ontology for linguistic concepts. Seattle WA: Proceedings of the *Knowledge Technologies Conference*, 12 August 2013. <http://staff.washington.edu/farrar/documents/inproceedings/FarLewLang02a.pdf>

Fillmore, C. 1992. 'Corpus linguistics' or 'computer-aided armchair linguistics'. In *Directions in Corpus Linguistics*, J. Svartik (ed.), 35–45. Berlin: Mouton de Gruyter.

Fishman, J.A. 1991. *Reversing Language Shift: Theory and Practice of Assistance to Threatened Languages*. Bristol: Multilingual Matters.

Fishman, J.A. (ed.). 2001. *Can Threatened Languages Be Saved? Reversing Language Shift, Revisited: A 21st Century Perspective*. Bristol: Multilingual Matters.

Garrett, E. 2014. Participant driven language archiving. In *Language Documentation and Description* 12, D. Nathan & P.K. Austin (eds), 68–84. London: SOAS.

Gawne, L., Kelly, B.F., Berez, A. & Heston, T. 2015. Putting practice into words: Fieldwork methodology in grammatical descriptions. Hawaii: Paper presented at *ICLDC 4 Conference*.

Grant, C. 2014. *Music Endangerment: How Language Maintenance Can Help*. Oxford: OUP. doi: 10.1093/acprof:oso/9780199352173.001.0001

Grenoble, L.A. 2010. Language documentation and field linguistics: The state of the field. In *Language Documentation: Practice and Values*, L.A. Grenoble & N. L. Furbee (eds), 289–309. Amsterdam: John Benjamins. doi:10.1075/z.158.28gre

Grenoble, L.A. 2011. Language ecology and endangerment. In *The Cambridge Handbook of Endangered Languages*, P.K. Austin & J. Sallabank (eds), 27–45. Cambridge: CUP. doi:10.1017/CBO9780511975981.002

Grenoble, L.A. & Whaley, L.J. 2006. *Saving Languages: An Introduction to Language Revitalization*. Cambridge: CUP.

Grinevald, C. 2003. Speakers and documentation of endangered languages. *Language Documentation and Description* 1: 52–72.

Gullberg, M. 2012. Bilingual multimodality and language documentation. In *Potentials of Language Documentation: Methods, Analyses, and Utilization* [Language Documentation and Conservation Special Publication 3], F. Seifart, G. Haig, N. P. Himmelmann, D. Jung, A. Margetts & P. Trilsbeek (eds), 46–53. Honolulu HI: University of Hawai'i Press.

Hale, K., Krauss, M., Watahomigie, L.J., Yamamoto, A.Y., Craig, C., Jeanne, L.M. & England, N.C. 1992. Endangered languages. *Language* 68(1): 1–42.

Harrison, K.D. 2005. Ethnographically informed language documentation. *Documentation and Description* 3, 22–41. London: SOAS,

Haugen, E. 1972. *The Ecology of Language*. Stanford CA: Stanford University Press.

Himmelmann, N.P. 1998. Documentary and descriptive linguistics. *Linguistics* 36: 161–195. doi:10.1515/ling.1998.36.1.161

Himmelmann, N.P. 2002. Documentary and descriptive linguistics. In *Lectures on Endangered Languages*, Vol. V, O. Sakiyama & F. Endo (eds), 37–83. Kyoto: Endangered Languages of the Pacific Rim,

Himmelmann, N.P. 2006. Language documentation: What is it and what is it good for? In *Essentials of Language Documentation* [Trends in Linguistics. Studies and Monographs 178], J. Gippert, N. P. Himmelmann & U. Mosel (eds), 1–30. Berlin: Mouton de Gruyter. doi:10.1515/9783110197730

Himmelmann, N.P. 2008. Reproduction and preservation of linguistic knowledge: Linguistics' response to language endangerment. *Annual Review of Anthropology* 37: 337–350. doi:10.1146/annurev.anthro.37.081407.085226

Himmelmann, N.P. 2012. Linguistic data types and the interface between language documentation and description. *Language Documentation and Conservation* 6: 187–207.

Hinton, L. 1993. *Flutes of Fire*. Berkeley CA: Heyday Books.

Hinton, L. 2002. *How to Keep your Language Alive*. Berkeley CA: Heyday Books

Hinton, L. 2011. Revitalization of endangered languages. In *The Cambridge Handbook of Endangered Languages*, P.K. Austin & J. Sallabank (eds), 291–311. Cambridge: CUP. doi:10.1017/CBO9780511975981.015

Hinton, L. 2013. *Bringing Our Languages Home: Language Revitalization for Families*. Berkeley CA: Heyday Books

Hinton, L. & Hale, K. (eds). 2001. *The Green Book of Language Revitalization in Practice*. San Diego CA: Academic Press. doi:10.1163/9789004261723

Holton, G. 2014. Mediating language documentation. In *Language Documentation and Description* 12, B. Nathan & P.K. Austin (eds), 37–52. London: SOAS.

Hornberger, N.H. 2010. Language shift and language revitalization. In *The Oxford Handbook of Applied Linguistics*, R.B. Kaplan (ed.), 365–373. Oxford: OUP.

Hymes, D. 1964. Introduction: Toward ethnographies of communication. *American Anthropologist* 66(6): 1–34. doi:10.1525/aa.1964.66.suppl_3.02a00010

Kapono, E. 1995. Hawaiian language revitalization and immersion education. *International Journal of the Sociology of Language* 112: 121. doi:10.1515/ijsl.1995.112.121

Kroskrity, P.V. 2015. Designing a dictionary for an endangered language community: Lexicographical deliberations, language ideological clarifications. *Language Documentation and Conservation* 9: 140–157.

Lehmann, C. 2001. Language documentation: A program. In *Aspects of Typology and Universals*, W. Bisang (ed.), 83–97. Berlin: Akademie Verlag.

Linn, M.-A. 2014. Living archives: A community-based language archive model. In *Language Documentation and Description* 12, D. Nathan & P.K. Austin (eds), 53–67. London: SOAS.

Lüpke, F. 2009. Data collection methods for field-based language documentation. In *Language Documentation and Description* 6, P.K. Austin (ed.), 53–100. London: SOAS.

Lüpke, F. & Storch, A. 2013. *Repertoires and Choices in African Languages*. Berlin: Mouton de Gruyter. doi:10.1515/9781614511946

Mayer, M. 1969. *Frog Where Are You?* New York NY: Dial Books for Young Readers.

Mühlhäusler, P. 1992. Preserving languages or language ecologies? A top-down approach to language survival. *Oceanic Linguistics* 31(2): 163–180. doi:10.2307/3623012

Mühlhäusler, P. 2000. Language planning and language ecology. *Current Issues in Language Planning* 1(3): 306–367. doi:10.1080/14664200008668011

Nathan, D. 2004. Documentary linguistics: Alarm bells and whistles? Seminar presentation, *SOAS*. 23 November 2004.

Nathan, D. 2006. *Sound and unsound documentation: Questions about the roles of audio in language documentation*. Washington DC: Paper presented at the Georgetown University Roundtable on Linguistics.

Nathan, D. 2009. The soundness of documentation: Towards an epistemology for audio in documentary linguistics. *Journal of the International Association of Sound Archives* 33. <www.iasa-web.org/book/iasa-journal-no-33-june-2009> (22 March 2012).

Nathan, D. 2010a. Sound and unsound practices in documentary linguistics: Towards an epistemology for audio. *Language Documentation and Description* 7, 1–17. London: SOAS.

Nathan, D. 2010b. Archives 2.0 for endangered languages: From disk space to MySpace. *International Journal of Humanities and Arts Computing* 4(1–2): 111–124.
doi:10.3366/ijhac.2011.0011

Nathan, D. 2014. Access and accessibility at ELAR, an archive for endangered languages documentation. In *Language Documentation and Description* 12, D. Nathan & P.K. Austin (eds), 187–208. London: SOAS.

Nathan, D. & Austin, P.K. 2004. Reconceiving metadata: Language documentation through thick and thin. In *Language Documentation and Description* 2, P.K. Austin (ed.), 179–187. London: SOAS.

Newman, P. 2003. The endangered languages issue as a hopeless cause. In *Language Death and Language Maintenance* [Current Issues in Linguistic Theory 240], M. Janse & S. Tol (eds), 1–13. Amsterdam: John Benjamins. doi:10.1075/cilt.240.03new

Newman, P. 2009. Fieldwork and fieldmethods in linguistics. *Language Documentation and Conservation* 3(1): 113–125.

Newman, P. 2013. The law of unintended consequences: How the endangered languages movement undermines field linguistics as a scientific enterprise. Seminar presented at *SOAS*, 15th October.

Niedzielski, H.Z. 1992. The Hawaiian model for the revitalization of native minority languages and cultures. In *Maintenance and Loss of Minority Languages* [Studies in Bilingualism 1], W. Fase, K. Jaspaert & S. Kroon (eds), 369–384. Amsterdam: John Benjamins. doi:10.1075/sibil.1.24nie

Penton, D., Bow, C., Bird, S. & Hughes, B. 2004. Towards a general model for linguistic paradigms. In *Proceedings of EMELD-04*, 12 August 2013. <http://emeld.org/workshop/2004/bird-paper.html>

Robins, R.H. & Uhlenbeck, E.M. 1991. *Endangered Languages*. New York NY: Berg.

Romaine, S. 2007. Preserving endangered languages. *Language and Linguistics Compass* 1(1–2): 115–132. doi:10.1111/j.1749-818X.2007.00004.x

Sallabank, J. 2012. From language documentation to language planning: Not necessarily a direct route. In *Potentials of Language Documentation: Methods, Analyses, and Utilization* [Language Documentation and Conservation Special Publication 3], F. Seifart, G. Haig, N.P. Himmelmann, D. Jung, A. Margetts & P. Trilsbeek (eds), 118–125. Honolulu HI: University of Hawai'i Press.

Sallabank, J. 2013. *Endangered Languages: Attitudes, Identities and Policies*. Cambridge: CUP. doi:10.1017/CBO9781139344166

Schultze-Berndt, E. 2006. Linguistic annotation. In *Essentials of Language Documentation*, J. Gippert, N.P. Himmelmann & U. Mosel (eds), 213–251. Berlin: Mouton de Gruyter.

Seifart, F. 2008. The representativeness of language documentations. *Language Documentation and Description* 5, 60–76. London: SOAS,

Sherzer, J. 1987. A discourse-centered approach to language and culture. *American Anthropologist* 89(2): 295–309. doi:10.1525/aa.1987.89.2.02a00010

Spolsky, B. 1989. Maori bilingual education and language revitalization. *Journal of Multilingual and Multicultural Development* 10(2): 89–106. doi:10.1080/01434632.1989.9994366

Spolsky, B. 2003. Reassessing Maori regeneration. *Language in Society* 32: 553–578. doi:10.1017/S0047404503324042

Sugita, Y. 2007. Language revitalization or language fossilization? Some suggestions for language documentation from the viewpoint of interactional linguistics. In *Proceedings of First Conference on Language Documentation and Linguistic Theory*, P.K. Austin, O. Bond & D. Nathan (eds), 243–250. London: SOAS.

Thieberger, N. 2012. Counting collections. *Paradisec Blog*, 29 November 2012. <http://www.paradisec.org.au/blog/2012/11/counting-collections/> (6 June 2015).

Thieberger, N., Margetts, A., Morey, S., Musgrave, S. & Schembri, A. 2012. Assessing curated corpora as research output. Paper presented to the *Annual Conference of the Australian Linguistic Society*, University of Western Australia.

Whalen, D. 2004. How the study of endangered languages will revolutionize linguistics. In *Linguistics Today – Facing a Greater Challenge*, P. van Sterkenburg (ed.), 321–342. Amsterdam: John Benjamins. doi:10.1075/z.126.16wha

Whalen, D. 2013. Response to Paul Newman. Blog post dated 3 December 2013. <http://elar-archive.org/blog/response-to-paul-newman-by-doug-whalen/> (27 March 2015).

Woodbury, T. 2003. Defining documentary linguistics. *Language Documentation and Description* 1, 35–51. London: SOAS.

Woodbury, T. 2011. Language documentation. In *The Cambridge Handbook of Endangered Languages*, P.K. Austin & J. Sallabank (eds), 159–186. Cambridge: CUP. doi:10.1017/CBO9780511975981.009

Woodbury, A.C. 2014. Archives and audiences: Toward making endangered language documentations people can read, use, understand, and admire. In *Language Documentation and Description, 12: Special Issue on Language Documentation and Archiving*, D. Nathan & P.K. Austin (eds), 19–36. London: SOAS.

Yamada, R.-M. 2007. Collaborative linguistic fieldwork: Practical application of the empowerment model. *Language Documentation and Conservation* 1(2): 257–282.

The brief existence of Saipan Carolinian

A study of a vanishing language storing valuable linguistic and historical insights on the tongue of its speakers

S. James Ellis
University of Hawaii

Saipan Carolinian emerged as a unique Micronesian language in its own right only about 100 years ago due to a series of migrations from a vast language continuum in the western Caroline Islands to an uninhabited island several hundred miles to the north, an island today known as Saipan. In view of the very shallow history of Saipan Carolinian the value in documenting this language is not for the preservation of ancient traditional knowledge – its parent languages, with their rich ancient traditions will, in fact, likely outlive this offspring language. Rather, it is Saipan Carolinian's inherited combination of linguistic features from its diverse but very-closely related parent languages that make it well-deserving to document. It offers valuable insights into how a blended language like Saipan Carolinian can derive phonemic features from one source language (Woleaian) and lexical features from another source language (Polowat) – two languages which are mutually unintelligible. This study also reveals that the history of migration events to Saipan is stored on the tongues of its speakers – a story which happens to be in contrast to what is believed by its speakers, but is nonetheless one that is historically accurate based on well-documented written records.

Keywords: Caroline Islands, documentation, lexical, Micronesia, migration, preservation, Saipan Carolinian, written record

1. Introduction

There are many urgent reasons for documenting threatened and endangered languages. Perhaps the most common one is summed up well in the Living Tongues mission statement, '*Every two weeks the last fluent speaker of a language passes*

on and with him/her goes literally hundreds of generations of traditional knowl-edge encoded in these ancestral tongues.'[1] Although Saipan Carolinian – an Aus-tronesian language spoken in the Northern Marianas Islands of Micronesia – is on track to become one of those disappearing tongues, it would not appear to warrant the time and effort it takes to document it; at least, not on the basis of the argument just given. For one thing, Saipan Carolinian has a very shallow his-tory – it emerged as a language in its own right only a century ago. Furthermore, all the languages that spawned the Saipan Carolinian language – each with their own ancient history – are still spoken today on the outer atolls distributed across the Caroline Island chain. In fact, the "traditional knowledge encoded" in those outer-island languages will most likely outlive the Saipan Carolinian language it-self. How the global community stands to benefit by documenting languages that have recently come into existence, due to the movement of people from their home areas, (including pidgins, creoles, mixed languages and so on) is not nec-essarily through the preservation of traditional knowledge. Rather, it stands to benefit from a fuller understanding of what can emerge linguistically and socially when various language or dialect communities come into long-term contact with each other. This case study of Saipan Carolinian provides us with further infor-mation on what can emerge when three separate (but closely related) language communities start a new community together on an uninhabited island … that being the island of Saipan.

This study will show how Saipan Carolinian conglomerated itself as a new language from many other Chuukic languages, primarily, three source languag-es – Woleaian (WOL), Satawalese (SAT), and Polowatese (POL) – each separated from the other by a hundred miles or more of deep ocean. Those three source lan-guages, however, did not contribute equally to the new language born on Saipan. While the three languages as well as Saipan Carolinian basically share a common grammar, the sound system of Saipan Carolinian is influenced more by WOL while the lexicon is influenced more by POL – two languages that are mutually unintelligible. That is to say, the marked phonemes of WOL have survived since the earliest years of Saipan Carolinian evolution even in the face of a greater pop-ulation influx of POL-speaking populations over the course of its evolution.

The study will also show that Saipan Carolinian's lexicon is an amalgamation of all three source languages. There are many instances where Saipan Carolinian has borrowed a word from each of its source languages in which all three source

---

1.   Living Tongues Institute for Endangered Languages <http://www.livingtongues.org/>.
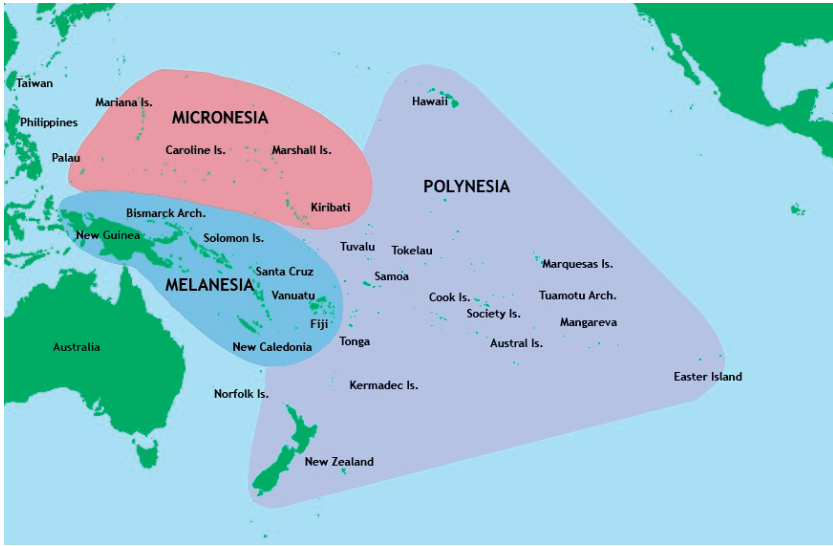
Figure 1. Pacific culture areas (Wikipedia Commons)

words share a common semantic load[2] – but with none being cognate with the other. Thus, the one united lexicon of Saipan Carolinian has developed a greater inventory of synonyms than any one of its founding languages.

The study will further demonstrate how the migration pattern hidden in the linguistic code of any fluent Saipan Carolinian speaker today betrays the common wisdom regarding the order in which the outer islanders settled Saipan; and how that spoken linguistic evidence is supported by historical documentation covering migrations to Saipan during the 1800's.

Documenting Saipan Carolinian yields the kind of linguistic findings that none of its source languages are able to offer. The fact that no language other than Saipan Carolinian has the capacity to tell this particular story validates the time and effort devoted to documenting it; and the same can be said for other languages in the world that have been newly created through the contact of their respective speakers. Each has the potential to add something new to our understanding of how language works.

---

2. My usage of the term "semantic load" refers to words in one language that correspond to words in another language in terms of their respective range of meaning.

1.1   Oral history[3]

Two hundred years ago a sailing canoe from Satawal, a tiny atoll in the string of the Caroline Islands of Micronesia, managed to navigate six hundred miles north over open ocean to a lush island supporting a small ridge of mountains and, of particular importance, a very large surrounding reef. It made a perfect destination to establish a new settlement, a refuge for outer islanders whose atolls had been destroyed by super-typhoons. The island was known to the navigators as Seypel[4] – today called Saipan, the capital of the Northern Mariana Islands – and it had lain uninhabited by people and free to grow wild for over a hundred years.

As described in both historical documentation and oral tradition, Chief Agharubw and his small party, having received permission from the Spanish administration on Guam, choose a coastal location on Saipan to establish a colony and build a permanent settlement.[5] Certainly, this would have included a traditional canoe house – an *utt*. The *utt* would not only provide shelter for canoes and supplies and a safe place to sleep, it would also mark with authority the establishment of a Carolinian settlement. That settlement was called Arabwal, currently called Garapan. Chief Agharubw and his group of Satawalese migrants along with Chief Nguschul from an atoll near Satawal would, over the next several years, be joined by a great number of canoes from a number of other atolls – primarily Woleai, Lamotrek, and other atolls in the region west of Satawal. Some fifty years later, waves of additional migrants would settle on Saipan from atolls to the east of Satawal – primarily, Polowat, Pulusuk, Pollap and Namonuito. Over the course of a hundred years the many different forms of related speech would blend into one, a new Carolinian language, today called Saipan Carolinian.

This information comes to us today not only from oral tradition (see, e.g., Farrell 1991). It can also be found among old Spanish records from the1800s (Driver & Brunal-Perry 1996). Interestingly, the main events of the Saipan migration can also be extracted from the linguistic features that make up the Saipan Carolinian language. These features will be investigated in Section 2 of this paper.

---

3.   My apologies to any Carolinian readers of this document for not including all the rich and specific details of Refaluwasch history. The intent of this paper is limited to describing the richness of the Saipan Carolinian language.

4.   P.c. the late Frank Olopai (ca. 1972).

5.   Historical information on the earliest Carolinian colony on Saipan is thoroughly covered in Ellis (2012: 295–302).

1.2   Saipan Carolinian (SpnCRL) language context

Saipan Carolinian (CRL or, more specifically, SpnCRL) is just one of two indigenous Carolinian languages in the Northern Mariana Islands (a commonwealth of the United States) together representing 2,000–3,000 people (Census Bureau 2002), living mainly on the island of Saipan (see Figure 2). Speakers of SpnCRL have traditionally populated the central and southern portion of Saipan while speakers of the other Carolinian language, Talaabog (TAL) (and identified as TPV in the Ethnologue) – a language most often referred to in previous literature as Northern Saipan Carolinian and associated with the three-letter (CRN) – have traditionally lived in the village of Tanapag located in the northern portion of Saipan (see Figure 3). It is important to note at this point that the TAL/CRN speech community has a very different history from that of the SpnCRL community (itself made up of three slightly different but sometimes socially important dialects). While the ancestors of TAL speakers migrated from essentially one source atoll, Namonuito (or, Nómwonweité, as pronounced locally (Odango 2015)), the SpnCRL speakers migrated largely from three separate source atolls, thus creating the interesting linguistic circumstances of this study. It is due to that history that the linguistic features of SpnCRL (Southern Saipan Carolinian) are the focus of this study – exclusive of TAL – although TAL will be mentioned from time to time during the course of this study.[6]

There is yet another indigenous language spoken on Saipan, called Chamorro, which is also the indigenous language of Guam (a territory of the United States). There are 10,000 Chamorro who share the island of Saipan with the Saipan Carolinians (PINA 2014) thus making Chamorro the more dominant indigenous language on Saipan. Saipan Carolinian, a nuclear Micronesian Oceanic language, is only remotely related to Chamorro which is not part of the Micronesian language family nor even the Oceanic language family. Oceanic is a branch of Austronesian[7] that includes Micronesia, Melanesian and Polynesian languages, as illustrated in Figure 1.

---

6.   During my work with the northern Carolinian community in Tanapag village from 1987–2005 I found there to be great confusion among Carolinian speakers in general as to whether the term "Enne" referred just to the Tanapag speakers or if it also referred to one of the three dialects of southern Saipan Carolinian: "Elle," Enne," and "Elle-Enne." CRN speakers decided to differentiate the northern language by referring to it as "Talaabog" (p.c. Sarapao, Ruak, etc.). Thus, I use the code (TAL) in my writings.

7.   The vast Austronesia language family covers both Western and Eastern Malayo-Polynesian from Madagascar to Easter Island. Oceanic is part of Eastern Malayo-Polynesian, whereas
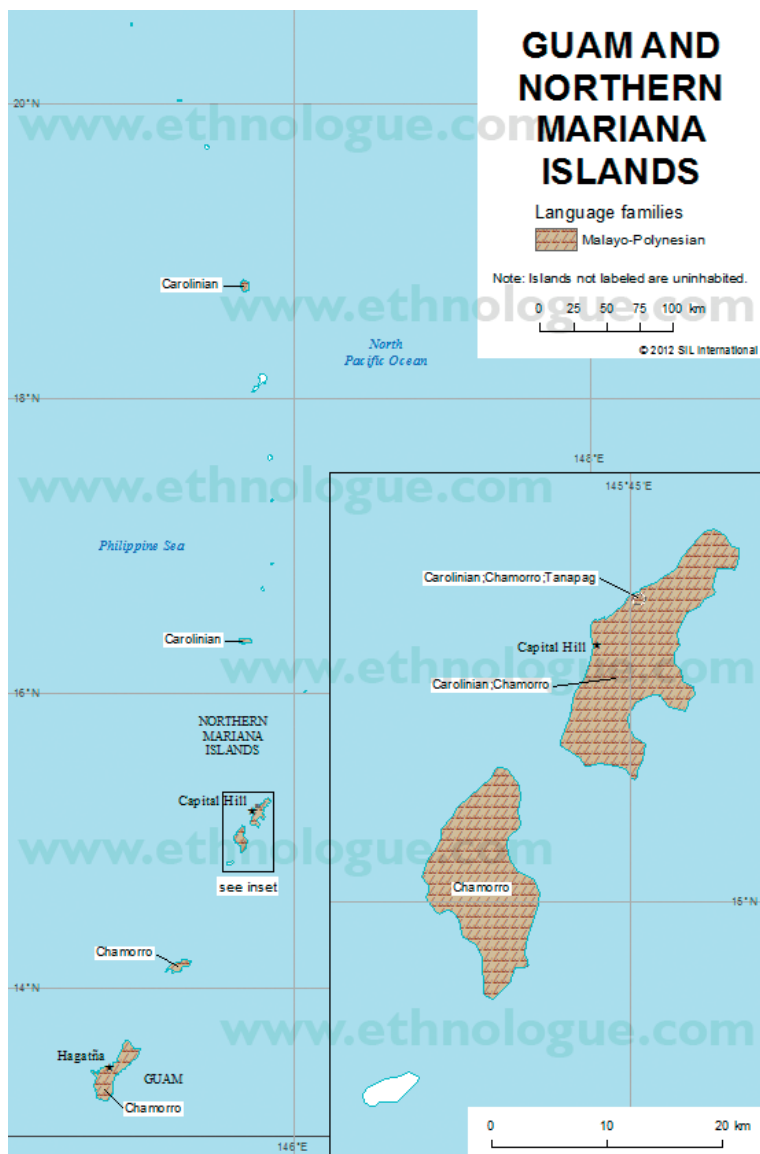
Figure 2. Mariana Islands chain (Lewis et al. 2014)

Western Malayo-Polynesian is composed of, in part, many Southeast Asian languages, most Philippine languages, and two languages in the Micronesian region: Chamorro and Paluan.
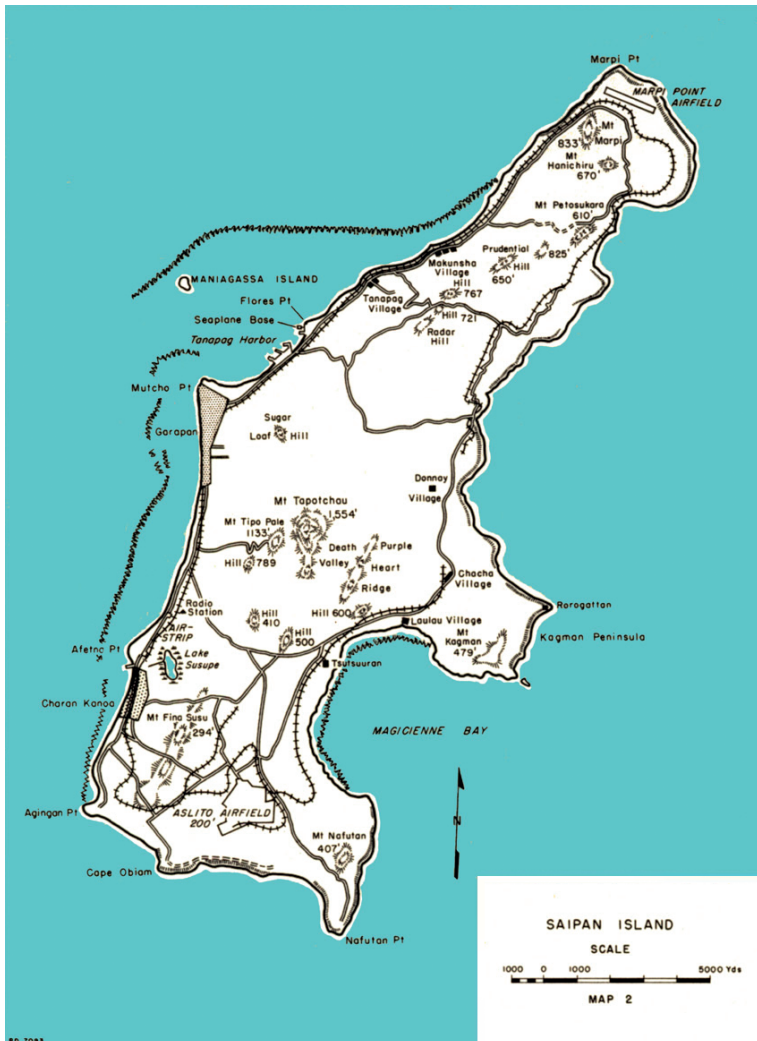
Figure 3. Saipan Island (Hoffman 1950)

The three-letter abbreviations for the Chuukic languages referred to in this paper are as follows.

> In the Republic of Palau:
> Tobian/Hatohobeian (TOB), Sonsorolese (SON),

> In the Federated States of Micronesia (FSM):
> Ulithian (ULI), Woleaian (WOL), Satawalese (SAT), Polowatese (POL), Pollapese (PLP), Namonuito/Nómwonweité (NAM), Mortlockese (MRT), Lagoon Chuukese (CHU),
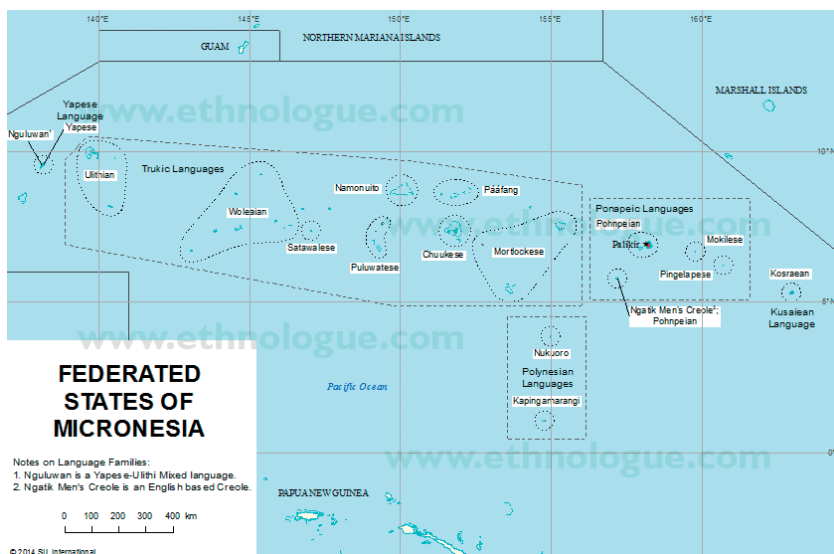
Figure 4. Chuuk (Trukic) languages (large circle) in the FSM (Lewis et al. 2014)

> Northern Mariana Islands:
>> Saipan Carolinian (SpnCRL), and Talaabog (TAL or CRN), located in
>> Tanapag, Saipan (see footnote 6 regarding the CRN nomenclature).

The chain of Carolinian languages from TOB, at the western edge, stretching 3,000 kilometers to CHU, at the eastern edge, are most typically referred to as Chuukic or Trukic languages. Since the focus of this paper is on the Carolinian language spoken on Saipan, the atoll communities and their languages will be referred to as 'outer-island Carolinian,' or, especially when including Saipan Carolinian, just the term Carolinian to apply to the whole continuum. As illustrated in Figure 4, each island or island group (within the large dotted circle marked as Trukic Languages) has a unique language, including SON and TOB in Palau and SpnCRL; and yet they share a very similar grammar, albeit with hundreds of small language-specific differences. In spite of that, only adjacent island groups are intelligible one to the other, for all practical purposes. For example, a Woleaian with only brief or perhaps no exposure to either Ulithian or Satawalese speakers will be able to understand either neighboring language when visiting those atolls. But the same Woleaian, though able to recognize many Polowat words that are cognate with Woleaian, would not be able to follow a conversation between two Polowatese speakers. And he would follow virtually nothing of Chuukese speech. This is because even though the grammars are so similar and even though there is a certain percentage of both phonemes and lexemes that all Carolinian languages share in
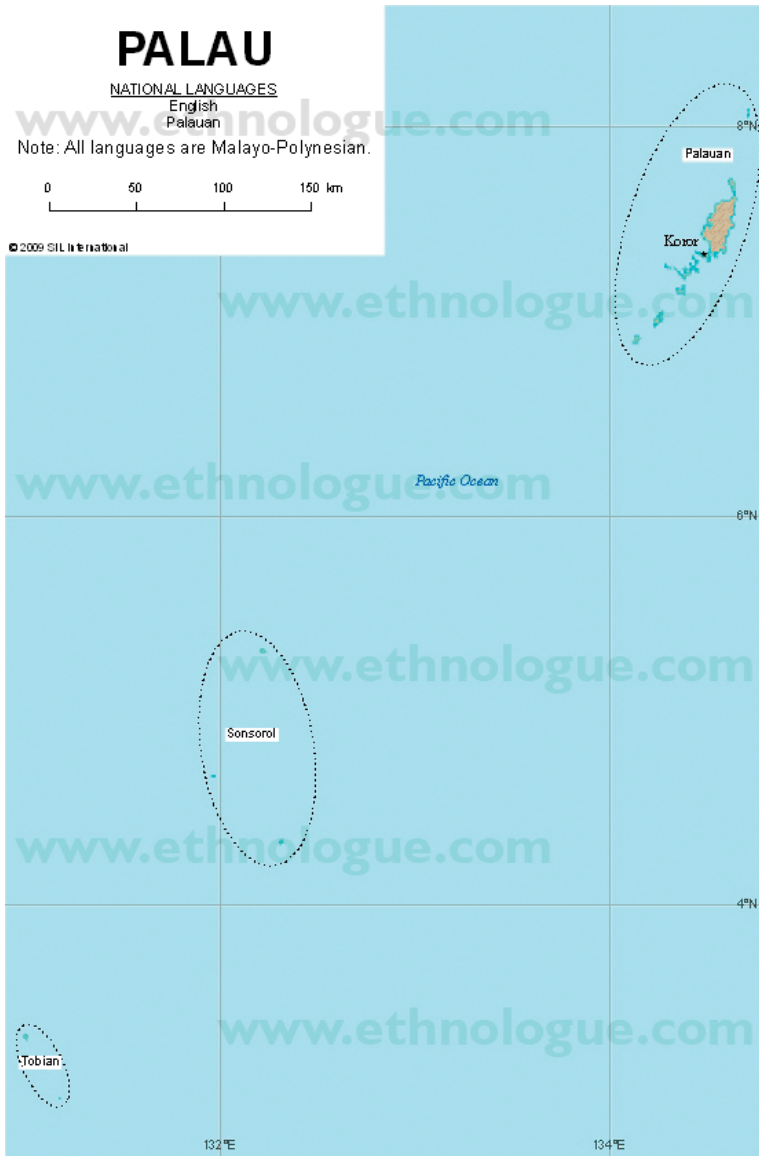
Figure 5. Sonsorol and Tobian, Chuukic languages in Palau (Lewis, M. Paul (ed.) 2009)

common, there are nonetheless significant differences between both phoneme and lexeme inventories – not to mention great differences in prosodic patterns – which create significant barriers to intelligibility.

The fact that there is from one to two hundred kilometers – and sometimes much more – of open ocean between island groups, with strong currents and
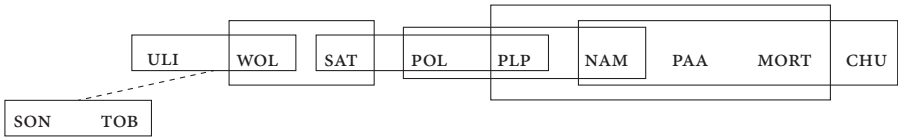
Figure 6. Language links within the Carolinian Language Continuum
(adapted from Quackenbush 1968)
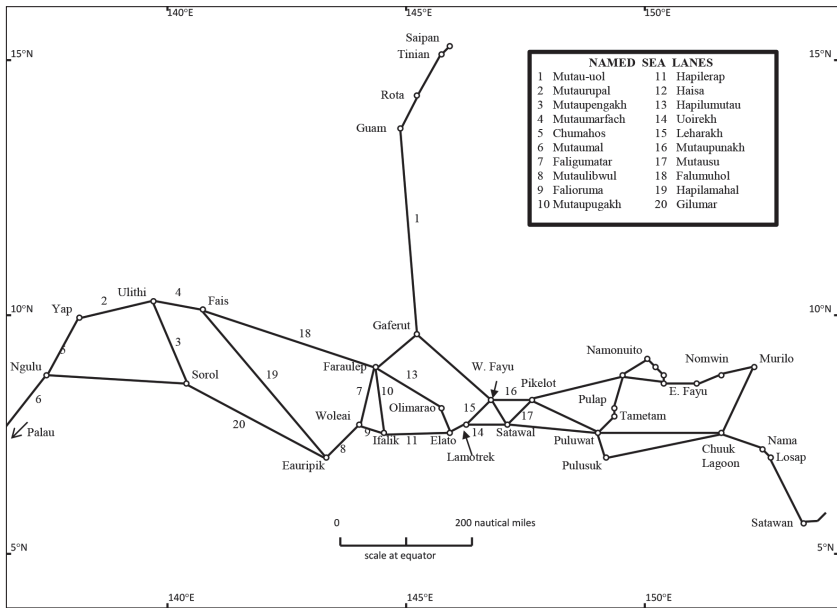
Figure 7. Sea lanes of the Western Carolines (adapted from D'Arcy 2006: 154)

periodic typhoons, helps to explain how remote the Carolinian speech com-
munities are from each other. On the other hand, even junior navigators were
qualified to sail the so-called "short stretches" of 100–200 kilometers between ad-
jacent islands (Marck 1986). And due to the technology of the Carolinian sailing
canoe and extreme knowledge of sky and ocean of the master navigators, they
were able to transport cargo and people anywhere in the Carolinian islands and
beyond – thus explaining the persistence of the language continuum. Routes be-
tween islands had unique names, as shown in Figure 7, and required a special set
of memorized instructions to apply during a journey (D'Arcy 2006).

There was also another linguistic phenomenon that allowed this Carolinian
network to persist; called "Language Bending" (refer to Ellis 2007 for a full de-
scription on this phenomenon). Language bending is a speaking skill that provides

successful communication between island groups whose respective languages are beyond the limits of "mutual intelligibility." Carolinians have the capacity to alter their own particular speech in a way that approximates, as nearly as they can, linguistic features that they know – or at least think they know – are used by speakers of a distant language group. Satawalese speakers, for example, can typically make regular sound-correspondence changes used by a Ulithi speaker, along with some changes in lexicon, word order, and particularly intonation, and thereby make themselves understood to Ulithians. And in a similar way, those Satawalese speakers can also alter their own speech to incorporate linguistic features which they expect a Chuukese speaker to use and are thus able to make themselves understood by the Chuukese speaker. Without language bending, attempts by typical SAT and CHU speakers to converse naturally with each other are futile (see Quackenbush 1968: 101–105). A small portion of one language bending example follows (adapted from Ellis 2007: 8):

(1)   Text translated into Ulithian:
Choako   re   yengaaeng yigla yathol   grade seg mo ruwow.
group.these 3PL work   today belong.of grade ten and two

(2)   Satawalese-bent-to-Ulithian speech:
Choka re engang ikila ngo oauton seik me ruou rak.

(3)   Original Satawalese text:
Rhoaka   re   engang ikina ie   oauton   seik me ruou rhak.
group.these 3PL work   now CONJ belong.of ten and two   grade

(4)   Satawalese-bent-to-Chuukese speech:
Choka repue angang ikina ie oauton engon me ruou chak.

(5)   Text translated into Chuukese:
Ekewe chóókka   ra   nnomw nóón engon me ruwuuw mwmwúch
the.PL group.these 3PL.PERF stay   inside ten and two   end
repwe   angaang.
3PL.ASP work

(6)   Free translation
"Only those12th graders were working now/today."

That is the linguistic landscape, then, that existed across the chain of Carolinian atoll languages at the outset of migration to Saipan just over 200 years ago. At that time there were not only no SpnCRL languages, there were also no people referred to as Saipan Carolinians. That would soon change, however, with the establishment of the first permanent Saipan Carolinian settlement in 1815.

1.3   Saipan Carolinian (SpnCRL) historical context

The Saipan Carolinian story starts early in the 1800s. It was during this time that a cycle of particularly severe typhoons was hitting the central Caroline atolls – which typically contained no more than one kilometer of dry land. When a severe typhoon passes over an atoll it can inundate the entire island and wash away everything, and everyone, that is not securely tied to a coconut tree.[8] All those that survive the storm are then faced with ruined crops and eventual starvation. It is only through the compassion and generosity of neighboring atolls that a population can recover.

As mentioned in Section 1.1, oral history tells us that during this time there was a particularly brave chief or two who managed to make it to Guam to ask the Spanish administration there for help. The Spanish were more than happy to have Carolinians living on Saipan. It was their mastery in sailing that brought about provision of a dependable source of meat and produce for the population on Guam (Arago 1823).

The Chamorro story on Saipan was very different. As a result of a long history of unfortunate events between the Spanish and the indigenous Chamorros, the Spanish gave up attempts to pacify the islands north of Guam and forcibly relocated the Chamorros to Guam where they could be more easily controlled. This happened in 1698 and resulted in the massive island of Saipan (in the view of atoll dwellers) being uninhabited from that time until 1815 when the Carolinians received permission from the Spanish to establish a new community there. These are the events that stimulated the series of migrations over a period of 100 years by a very diverse outer-island population.

Now the question arises, what kind of speech eventually evolved in the Saipan community in view of the fact that there were speakers from different language locations. Certainly language bending would have been important in the earlier days. But one would have thought that in a generation or two the wide range of languages heard on Saipan would have begun to gel into a common language. That, however, cannot be deduced from the historical record. What can be deduced from the historical record is that the language ecology was in a constant state of flux for most of the 1800s. Every time a few canoe loads of new migrants landed on Saipan it would have had an impact on the language ecology there. Added to that equation was the fact that the movement of Carolinians was not just one way. Any canoe heading back to the outer island atolls from Saipan gave

---

8.   An interesting point is made by a reviewer of this chapter that helps to explain how devastated atoll-dwellers survived, "The survival of their canoes at such times was and is accomplished to some extent by burying them when especially large storms were apparently developing."

Carolinians living there an opportunity to spend a year or more with relatives back home. This cyclical migration was a continuous process of movement both ways,[9] constantly moving populations back and forth.

The fluid composition of language variation between Saipan and the outer islands came to an abrupt stop when the Spanish gave up their Micronesia colony at the turn of the century and the succeeding colonial power, Germany, was enticed into making Micronesia their responsibility, in 1899. In the first decade of the 1900s the German administration, with headquarters on Saipan, would put a stop to transportation by sailing canoe and the succeeding Japanese administration would continue the policy (McCoy 1976: 356–357) – too dangerous, too much loss of life. So, of course, the continuous cycles of language impact from the outer islands stopped. And shortly thereafter the wide range of speech variation on Saipan finally began to gel into one shared language. Children and grandchildren of the migrants began speaking a common language (SpnCRL) – albeit one littered with dialectal variation (as persists today to a limited extent). So it was not until perhaps around 1915 that the speech on Saipan finished its decades of tumultuous gestation and could then be considered a new and unique Chuukic language in its own right within the Carolinian continuum. The language came to be known as Saipan Carolinian, or, in the Carolinian language, Refaluwasch ("people of our island").

Over the decades the Saipan Carolinian community essentially lost contact with their home lands and languages. Nonetheless, they did continue living in traditional ways as much as possible.

1.4   Saipan Carolinian (SpnCRL) intelligibility of other CRL languages

Most Saipan Carolinians today will say that SAT is easiest to understand. But my own language intelligibility research throughout the continuum (see Table 1) indicates, on the contrary, that they understand the POL speech better than SAT speech. This would seem most clearly due to two things. First, as mentioned, the prosody of SAT probably makes it sound more similar to SpnCRL than the

---

9.   One may wonder if the reverse migration of individuals back to their home atoll had the effect of increased language leveling within the greater Carolinian language continuum. Most likely so, since during the last 2000 years there has been a vast amount of lexical diffusion among all the languages in the Carolinian continuum originating from proto-Chuukic (e.g., Jackson 1983). At the same time, this diffusion of lexicon (and phoneme inventory) did not make all languages mutually intelligible. It is clear from historical records (especially from studies carried out during the 1908–1910 Südsee Expeditions (e.g., Krämer 1937)) that the picture of language intelligibility across the atolls was similar to what it is today.

Table 1. Saipan Carolinian and Talaabog language intelligibility test results (Ellis 2012)

Languages tested	Ulithian text	Woleaian text	Satawalese text	Pollapese text	Chuukese text
Respondents from:					
SpnCRL (1989 testing)	4%	11%	48%	88%	5%
TAL (1993 testing)*	–	–	67%	94%	90%

(Also from 1993 testing, the following:)

Namonuito respondents scored 47% with Saipan Carolinian (SpnCRL) and 90% with Talaabog (TAL).

* Due to the level of interaction between SpnCRL and TAL speakers over a great many years any attempt at intelligibility testing would be highly skewed, thus, no testing was carried out between the two. These percentages also corroborate a comment given earlier which suggested that the TAL language has retained a more conservative form of speech than has SpnCRL.

other languages; although no objective study has yet been carried out to determine the degree to which that is true. Second, the Satawalese are masters at language bending, as discussed earlier in this section. They have the ability to hone in on the language of the one they are speaking to and to alter their own SAT speech to an extent that the hearer can understand what the Satawalese is saying. Because of that, the respective hearers assume SAT speech to be much closer to their own form of speech than is actually the case.

Table 1 shows the intelligibility-testing results for the Saipan Carolinians (Ellis 2012). Also included are some results for respondents from the village of Tanapag whose language, Talaabog (TAL), proved to retain so much of its source language, Namonuito, that it can be considered a dialect of that source language.[10] During my testing it became clear that when a group of speakers from a given atoll scored about 85% or *higher* with another Carolinian variety of speech, one with which they had not had significant previous contact, that indicated that the former variety of speech could be considered a dialect of the latter variety of speech. If, on the other hand, speakers from a given atoll scored about 85% *or lower* on the target language, the former variety of speech could not be consider a dialect of the latter, rather, the former must be considered a separate language from the latter (Ellis 2007: 13). This threshold figure of 85% – or very close to it – is commonly used as the threshold between language and dialect in other intelligibility studies worldwide (see e.g., Grimes 1992: 32 and 1995: 22; Casad 1974: 83–86).

---

10. A reviewer correctly pointed out that even as late as the 20th century the TAL/CRN language would still not have been considered intelligible to SpnCRL children from central and southern villages.

2.   Linguistic characteristics of SpnCRL, a blended Carolinian language

Since SpnCRL emerged directly from and only from other languages in the Carolinian continuum (with the exception of a number of loan words from colonial languages), the challenge has been to discover those points within the continuum from which linguistic features of SpnCRL were derived. In this section evidence is provided to show that WOL, SAT, and POL were the *primary* source languages of SpnCRL. It will be shown that SpnCRL did not inherit the vast majority of its linguistic DNA from just one of the three source languages – and certainly not SAT, the language in the middle, as one might expect and as so many believe. What did develop is that the new language ended up being an *amalgamation* of all three of its primary source languages (as will be explained in Section 2.3). This is interesting because it provides another example of the variety of outcomes that can occur when two or more language and/or dialect communities become combined together at a new geographic and social setting.

Many examples of the fusion[11] of languages/dialects in other contact situations were also presented at the 36th LAUD Symposium, including a report on Aragon, which is in a diglossic relationship with Spanish (see Abstract Hijazo-Gascón and Ibarretxe-Antunano), on Portuguese creole situations (see Abstract Nunes and Abstract Lee), on some 150 modern-day dialects of Aramaic, an ancient written language (see Abstract Khan), and so on. The diverse outcomes of a wide range of previously studied language and dialect contact situations have been well covered by, for example, Weinreich (1953) and Trudgill (1986). A particularly interesting domain of language variation under study in recent decades is the phenomenon of "mixed languages." A "mixed language" is one that is sourced from unrelated languages. The number of documented cases of "mixed languages" has continued to slowly increase since the seminal study on Michif, a language mixture of French and Cree.[12] But Saipan Carolinian is neither a Mixed Language nor is it a product of multiple dialects of a single language. Rather, it is a blend of several source languages in the center of the Carolinian continuum that are all closely related but are also clearly distinct.

To promote categorizing SpnCRL as something other than a "mixed language," a creole, a pidgin, or as a merger of dialects from one language, SpnCRL is

---

11. See Auer (1998) for a restricted definition of "Fused Lects."

12. In the Michif case the nouns tended to be absorbed from the one language and the verbs from the other (Bakker 1997).

referred to in this and other writings as a "blended language."[13] So now the question is, how is Saipan Carolinian structurally and lexically blended from three very closely-related but different source languages?

2.1 Lexical inventory of SpnCRL

As mentioned earlier, the SpnCRL grammar, particularly the morphology and syntax, is very similar to that of its three source languages, which is no surprise since the three, to a large extent, share a common grammar among themselves. It is also true that SpnCRL has inherited a very large portion of its lexicon from a common pool of lexemes shared not by just the three, but by all the other languages in the whole Carolinian continuum. As established by Quackenbush (1968), using a 585-word list built from a variety of sources including a modified version of the Swadesh 200-word list, and Jackson (1983), using a 200-word list of basic vocabulary designed specifically for Malayo-Polynesian languages (Jackson 1983:273), all members of the Chuukic language family – which constitute the whole of the Carolinian language continuum – share close to 50% of their core lexicons in common. This is illustrated in Table 2 where the quantity and percentage of lexical similarity between SpnCRL and each of the other Carolinian languages are shown. The source used for calculating the following figures comes from the first 300 words of the Quackenbush (1968) wordlist.

First, a note about word frequency. Using only the criterion of word similarity as the indicator of language similarity can give a false picture in language typology studies. This can especially be true in studies using large lexicons because words of high frequency carry greater weight in judging similarity between languages than do words of low frequency.[14] Standardized word lists used to de-

---

13. It is not clear to me if a universal acceptance of the terms 'pidgin' and 'creole' exists such that any given speech variety can be clearly classified as 'pidgin/creole' or not. Concepts of 'superstrate' (a dominat*ing* language like English) and 'substrate' (the dominat*ed* languages like indigenous Papuan or Melanesian languages), at least, do not apply to languages in the Carolinian continuum because no superstrate language has participated in the continuum and no substrate language exists there either; each is composed of a rich indigenous grammar that is very similar among all Carolinian languages including those on Saipan, which are of equal status to the other members of the continuum. A language like SpnCRL does not conform to pidgin, creole, or any other well-known nomenclatures. Thus, the term "blended language" is used to categorize its distinctive language/dialect relationships.

14. These comments on word frequency here are prompted by a reviewer's note that frequency of word usage – not just quantity of similarity – must be part of the calculus in judging linguistic typology. The reviewer writes that the "English language lexicon is composed roughly of one third Germanic, one third Latin and one third French lexemes, but the first 100 most frequent

Table 2. Percentages of similarities between SpnCRL and other languages (using first 300 items of the Quackenbush (1968) wordlist) (Ellis 2012)

Sonsorol	142.5	48%
Ulithi	160.0	54%
Woleai	213.0	71%
Satawal	237.0	79%
Polowat	227.5	76%
Pollap	187.5	63%
Namonuito	194.0	65%
Mortlocks	180.0	60%
Chuuk	180.0	60%

Profile of outer-island lexical similarity with SpnCRL
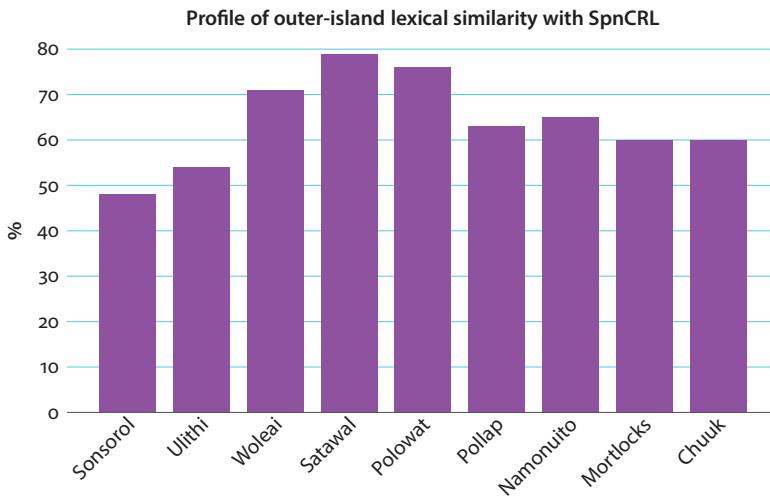
termine common heritage as used by Quackenbush and Jackson, however, are not selected at random but are, rather, selected on their likelihood of being stable over time. Such words are typically high frequency because they include common body parts, familial relationships, objects and activities encountered in daily life, and so on. The first 300 words in the Quackenbush study, which are the basis of this study, are all words of that nature.

Essential to the purposes of this study, however, are not the words of high frequency, per se. It is the words that are shared exclusively between SpnCRL and just one other of its source languages. Such words offer an indication that that particular word in the SpnCRL lexicon is sourced from the one respective source

words in that language are 90%+ Germanic in origin" thus implying, I assume, that German is more similar to English than the other two in spite of their equal amount of contribution.

language that has the same word in its lexicon – to the exclusion of that word in lexicons of any other source language. In fact, words of low frequency could be most useful for the purposes of this study since such words can be less widely known and more locally specialised, and thus, by extension, help to more securely establish the geographical source of human migrations to Saipan.

To make the point clear, Table 2 above displays the similarity between the 300 words in the SpnCRL lexicon, on the one hand, and the same words in the lexicons of other Carolinian languages, regardless of whether the match is exclusive or not.

Table 2 results suggest that SpnCRL has a lot in common with each of the Carolinian languages listed. However, to repeat, the high percentages of similarity are not necessarily a good indicator of which languages are the primary source languages of SpnCRL. What is particularly helpful to this study is finding those lexical items that SpnCRL shares with just one other Carolinian language since such exclusively-shared lexical items serve best to show which one or ones of potential source languages contributed a given lexical item to SpnCRL. Table 3 provides a few examples of exclusively-shared lexemes – which are also referred to here as form-meaning matches.

As mentioned above, my database of lexical form-meaning matches existing exclusively between SpnCRL and a given outer-island Carolinian language was built upon the first 300 words of the Quackenbush (1968) wordlist. Taking each of those 300 words, one word at a time, the SpnCRL lexicon was then searched thoroughly for any lexical items that matched any of the 300 words in meaning, not necessarily in form. Often there were two or more SpnCRL words that matched

Table 3. Lexical matches of both form and meaning shared exclusively by SpnCRL and a single other language

SpnCRL		WOL, SAT, POL exclusively-shared lexemes		
ppwatúr	"leprosy"	WOL	*bbatiur*	"leprosy"
schiyor	"choke"	WOL	*shiyor*	"be stuck in the throat"
yuumi	"bow and arrow"	SAT	*yumi*	"arrow"
olighát	"group of children" or occasionally, "child"	SAT	*wonikaet*	"child"
alúl	"color"	POL	*yanúyan*	"color"
		Examples from other outer-island languages		
fischi	"shoot at (something)"	PLP	*firhi*	"shoot"
wetil	"break/crack"	CHU	*wetin*	"be broken, shattered"
uleey	"slice it"	CHU	*wuneey*	"slice (it)"
peyas	"the dust from which you came"	MRT	*payas*	"ashes" (partial match semantically)

a given word in at least one of the outer-island Carolinian languages. (For instance, from Table 8 below: one word for "hole" in SpnCRL is *liibw*. SpnCRL has three other words – synonyms – that mean essentially the same thing: *ngaat, laas, and pwang*. Dictionaries for five of the languages in the continuum also have a form-meaning entry that matches the SpnCRL *ngaat*. Only three have a match for *laas* and only two have a clear match for *pwang*.) Following that procedure yielded a total of some 550 lexical entries in SpnCRL that could be used as a database to search for exclusive matches between SpnCRL and any other language.

Starting with this 550 word database, all available dictionaries and wordlists in each of the outer-island Carolinian languages were then mined for lexical items that matched – in both form and meaning – any of the 550 words in the SpnCRL dictionary. Surprisingly, only forty-nine word-meaning exclusive matches were found, as displayed in Table 4. Note that only a half point was given for cases where there was not a strong semantic match. An example of that is the MRT word *payas* corresponding to the SpnCRL word *peyas* in Table 3. The MRT sense of the word, "ashes," does not seem at first to have much in common with SpnCRL, "the dust from which you came." The Saipan Carolinians all became Christians, however, and at funerals the common phrase, "…ashes to ashes, dust to dust" could have been captured semantically by using the SpnCRL word *peyas* – thus the ½ point allotted to MRT in Table 3.

Following in Table 4 is the distribution of the 49 exclusive form-meaning matches. It is the profile of the chart in Table 4 that suggests WOL, SAT, and POL to be good candidates for being primary source languages of SpnCRL. In contrast, in the case of SON, for example, of the forty-nine exclusive matches only one word in the lexicon of Sonsorolese migrants – a word not used by Saipan migrants from other languages – made it into the permanent SpnCRL lexicon. It's not likely, therefore, that many Sonsorolese became residents of Saipan during the century of migration – for if so, one would expect them to have made more of a linguistic impact. At the other end of the scale, seventeen words in the lexicon of Polowatese migrates made it into the SpnCRL lexicon (as per Jackson & Marck 1991; Ellis & Fruit 2003–current). This would be unlikely to happen if, for instance, there were only intermittent Polowatese visitors to Saipan during the migration years – especially since the Polowatese were not high in the social scale (Hunter-Anderson & Zan 1996).

(The profile of the chart in Table 4 gives the impression that Chuukese stands out as another possible source language to SpnCRL. The Chuukese figure is misleading, however, because there are still no available dictionaries or comprehensive word lists devoted specifically for the languages of Pollap, Namonuito, or the Mortlocks. That Chuukese did not have a significant impact on Saipan speech during the years of migration – although there are a number of CHU lexical items

Table 4. Number and percentage of exclusively-shared form-meaning lexical matches

Language	Quantity of matches	Percentage of matches
Sonsorol	1	0.2%
Ulithi	2½	0.5%
Woleai	10	1.8%
Satawal	8½	1.5%
Polowat	17	3.1%
Pollap	2	0.4%
Namonuito	0	0.0%
Mortlockese	½	0.1%
Chuuk	7½	1.4%

Percentage of unique matches with SpnCRL for each language
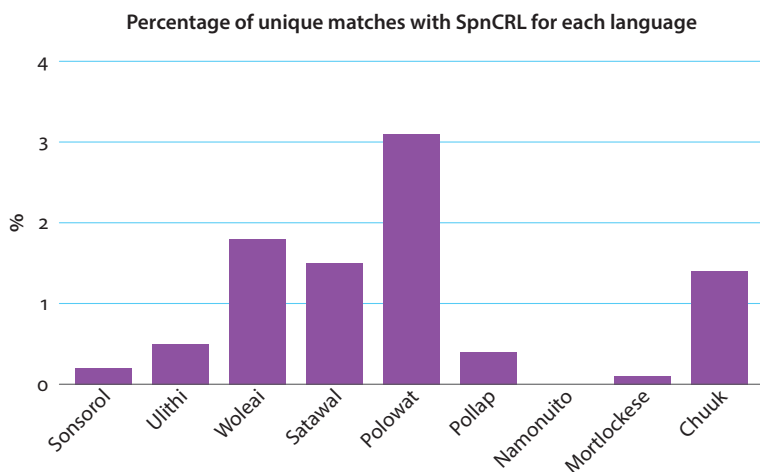
in SpnCRL's lexicon today – is bolstered by the fact that the Chuukese had no motivation to migrate to Saipan during the 1800s since the Chuuk Lagoon is mostly composed of high islands, thus providing a great deal of protection from typhoon swells. Additionally, as Quackenbush points out (1968: 86), "The figure of forty-nine exclusively-shared items for [or among all Carolinian/Trukic] languages is probably inflated by relatively recent diffusion outward from Truk…" thus giving CHU higher percentages than expected.)

With WOL, SAT, and POL being the most obvious source candidates for SpnCRL, the 550 form-meaning lexemes were again used to determine the number of SpnCRL's exclusively-shared lexemes pertaining *only* to WOL, SAT, and POL (excluding data from all other languages). The results generated Table 5, which shows clearly that POL has significantly more exclusively-shared form and meaning matches with SpnCRL than do the other two. And that is the basis for

Table 5. Form and meaning lexical-match data involving WOL, SAT, POL languages only

Languages	Matches with SpnCRL	
	Quantity	Percentage
only WOL matches SpnCRL	31.5	5.7%
only SAT matches SpnCRL	17.5	3.2%
only POL matches SpnCRL	60	10.9%

Percentage of unique form-meaning matches with SpnCRL for each language (or language pair) considering only Woleai, Satawal, and Polowat languages
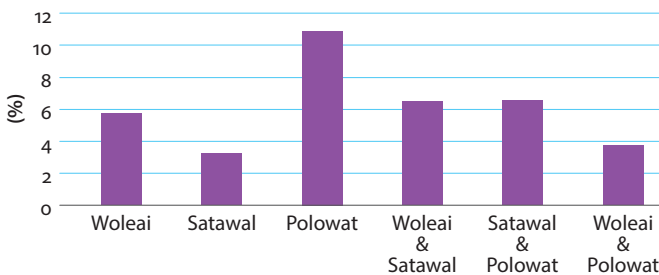
claiming that the Polowatese migrants were the most likely candidates as primary contributors to SpnCRL's current lexicon. It is also noteworthy that SAT, the candidate in the middle of the other two, has had the least impact on the SpnCRL lexicon. This would indicate that the Satawalese migrants contributed least to the SpnCRL language that was to emerge early in the 1900s.

Conclusions that can be drawn from Tables 4 and 5 stand in stark contrast to the overall consensus of Saipan Carolinians and others living on Saipan, that mistaken consensus being that the main population of southern Saipan Carolinians on Saipan are the descendants primarily of Satawalese migrants.

2.2   Phoneme inventory of SpnCRL

In addition to lexical inventories, phoneme inventories are a good place to look for differences among closely-related languages. So in view of the very lopsided lexical affiliations that align SpnCRL most with POL, one might expect that the Saipan Carolinian *phoneme* inventory would also be derived mostly from Polowat, or, again, at least the language in the middle, Satawal. But that is not the case.

The majority of phonemes in the inventories of WOL, SAT, and POL are drawn from a common pool of phonemes shared by all three. Accordingly, the

majority of SpnCRL phonemes have also been drawn from this same common pool. But there is a certain class of consonantal phonemes that enable a clear distinction between the languages to be drawn. It is this set of phonemes that Carolinians invariably use to demonstrate the differences between source languages (Ellis 2012). Asking a Saipan Carolinian what distinguishes his language from Satawalese or Polowatese almost always brings the response, "we say *sh* and they say *rh*" (see example for "shoot" in Table 3). Or, "we say *s* and they say *h*." These phonemes have resisted being used in free variation with each other. Carolinian languages consistently use one or the other to reflect their respective proto-Chuukic sound correspondent.

Although, the above is not true of the 'l' and 'n' distinction where there is great variation among the source languages. There is clearly free variation between [l] and [n] in SAT. POL reflects mostly [n] (see example for "color" in Table 3) but also [l], depending on the proto-Chuukic sound correspondent. WOL has merged both correspondents to [ɾ]. And SpnCRL (pertaining to the dominant southern "Elle" dialect) has merged both correspondents to [l]. It should also be noted at this point that the vast range of variation in the pronunciation of vowels in Carolinian languages renders them poor indicators of their affiliation (at least up to this stage in my research).

Apart from the proto-Chuukic [l] and [n] the remaining thirteen of the fifteen proto-Chuukic consonants can be used to establish both similarities and differences among all the Carolinian daughter languages. First, there are seven sound correspondences that have almost without exception retained their proto-Chuukic form, and WOL, SAT, and POL share completely in these phonetic

Table 6. IPA Chart (version: 2.1, released 2008, from www.sil.org/computing/ipahelp)

(PULMONIC)	Bilabial	Labioden	Dental	Alveolar	Postalv	Retroflex	Palatal	Velar	Uvular	Pharyngl	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k g	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r					ʀ		
Tap or Flap				ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lat. Fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lat. Approx.				l		ɭ	ʎ	ʟ			

┌ (NON-PULMONIC)			┌ OTHER SYMBOLS		
Clicks	Voiced implosives	Ejectives	ʍ Voiceless labial-velar fric.	ɕ ʑ Alveolo-palatal fricatives	
ʘ Bilabial	ɓ Bilabial	' as in:	w Voiced labial-velar approx.	ɺ Alveolar lateral flap	
ǀ Dental	ɗ Dntl/alvlr	p' Bilabial	ɥ Voiced labial-palatal approx.	ɧ Simultaneous ʃ and x	
ǃ (Post) alvlr	ʄ Palatal	t' Dntl/alvlr	ʜ Voiceless epiglottal fricative	Affricates and double articulation can be represented by 2 symbols joined by a tie bar if necessary.	
ǂ Palatoalvlr	ɠ Velar	k' Velar	ʢ Voiced epiglottal fricative		
ǁ Alvlr lateral	ʛ Uvular	s' Alveolar fricative	ʡ Epiglottal plosive	͡ or ͜ k͡p t͡s	

Table 7. Regular sound correspondences for SpnCRL and its source languages

WOL	SAT	POL	SpnCRL
[ʂ]	[ɻ]	[ɻ]	[**ʂ**]
[**x**]	[k]	[k]	[**x**]
[t͡ʂʲ]	[t͡ʃ]	[t͡ʃ]	[**t͡ʂʲ**]
[s]	[s]	[h]	[**s**]
[ɸʷ]	[pʷ]	[pʷ]	[bʷ]
[ɻ]	[**r**]	[**r**]	[**r**]

retentions. The seven are [p, f, t, m, mʷ, ng, w].[15] That leaves six corresponding phonemes that are useful for showing the picture of affinity that SpnCRL has with WOL, SAT, and POL. The first three rows of the six sound correspondences listed in Table 7 show that WOL and SpnCRL exclusively share the phonetic expression of these three phonemic sound correspondences. The table also shows that POL alone is dissimilar to SpnCRL in the fourth sound correspondent row. The fifth row shows a case where the SpnCRL correspondent is not phonetically equal to the WOL correspondant; nor is it so with SAT or POL, either. The last row shows the one instance where the SpnCRL correspondent shares the same phonetic features with both SAT and POL, but not with WOL. As a total picture, however, Table 7 makes it clear that SpnCRL has the highest phonemic affinity with WOL, and the lowest with POL. (In Table 7, SpnCRL phonemes that match one of the other language's phonemes are each marked in red and bold font.)

What can be inferred from the data in Table 7 is that the phoneme inventory of the WOL-speaking migrants – which includes the relatively large number of migrants from Lamotrek – sustained the greatest impact on what would become SpnCRL's phoneme inventory. Possible sociolinguistic reasons as to why the evolution of SpnCRL's *phoneme* inventory has followed such a different path through time from that of its *lexical* inventory will be taken up in Section 2.3 below.

2.3   Lexical amalgamation

The lexicon of each Carolinian language in the continuum is made up of a distinctive inventory of form-meaning correspondences. And while, as indicated earlier, all languages in the Carolinian continuum share roughly 50% of their lexical inventories, that implies that there is another 50% of their lexicons where there is a full range of differences among the languages. Of the three examples in Table 8,

---

15. The mʷ is a velarized consonant (back contraction of the tongue). Refer to the IPA Chart in Table 6 for pronunciation guides to the phonemes discussed in this section.

Table 8. Examples of SpnCRL lexical amalgamation (Ellis 2012)

form-meaning sets utilized across continuum	SON**	ULI	WOL	SAT	POL	PLP	NAM	MRT	TRK	SpnCRL
One f-m set in common for "skin"	giirI/ giin	giil	giil (g = [x])	giil	kiil	kiil	kiil	kiil	siin	ghiil (gh = [x])
Two different f-m sets for "belch"	gyrélY/ gygég	kylér/ kulér	xurér	kyrér	kyrér			kyrér		ghérér
						maky	maky		méky	maghú
Four different f-m sets for "hole"	riibwA	liibw	liibw	niipw						liibw
					naah	naah			naas	laas
			ngat (biyangi)	ngaat	ngaat		ngaat		ngaat	ngaat
								pwaang	pwaang	pwang

** SON and TOB forms separated by "/".

the first is the only one where all the languages share a single form-meaning correspondence in common. The Carolinian words for 'skin' are all cognate lexemes right across the whole continuum. The second example, 'belch,' however, is a case where two different lexical forms are used among the continuum members to capture the same meaning. There is therefore not one form-meaning set but two. And in the third example, 'hole,' while each lexical form carries the same meaning, there are nonetheless four different lexical forms used among the continuum that are used to capture the one meaning in common; thus, there are four form-meaning sets for 'hole.'

Note that in the second example in Table 8 the speakers of each Carolinian language use either the one lexical form or the other lexical form, not both; except for SpnCRL. SpnCRL speakers use both. And note in the third example that speakers in all the language communities, except CHU, use only one or two different lexical forms for the word 'hole.' SpnCRL speakers, however, use all four lexemes in their lexical inventory for the word 'hole.' These examples illustrate that there was a tendency among the early SpnCRL speakers to use all or many of the available lexical forms used during the era of migration to Saipan. In other words, the SpnCRL lexicon became an amalgamation, to some degree, of the corresponding lexical forms from the available source-language lexicons – at least, certainly more so than other Carolinian languages.

This capacity of Saipan Carolinian speakers to amalgamate their lexicon from the many languages of origin was something that Georg Fritz noticed (1911:7)

during his administration on Saipan during the 1899–1914 German era. He wrote, "*The difference in origin also explains the wealth of words related in meaning, which in turn makes the language of the Saipan Carolinians a more suitable means of communication than one of the other Carolinian dialects.*"[16]

What Fritz observed was that those Carolinian islanders on Saipan who were descendants of the early migrants to Saipan – the early migrants originating from throughout the Carolinian continuum – controlled more alternate forms of a given lexeme than did the Carolinians of other atolls in the continuum.

An earlier writer (Spoehr 1954) interpreted Fritz as follows:

> Fritz (1911, 7) noted that the Saipan Carolinian language contained a wealth of synonymous words and found that Saipan Carolinian interpreters, using the Saipan dialect, were possessed of a better medium of communication in the atolls west of Truk than were individuals from single atolls in this group.

The information in this section suggests that the SpnCRL lexicon has amalgamated a significant portion of its lexicon from all three of its source languages – and from additional languages in the continuum as well.

3.   Migration pattern reflected in SpnCRL speech

In addition to illustrating the phenomena of 'language bending' and 'amalgamation' the composition of the SpnCRL linguistic code is useful in yet another intriguing way. It contains a basic historical record of migration to Saipan from the outer islands during the 1800s.

The following section explores the documented historical record. How and why the linguistic code of any fluent SpnCRL speaker corroborates the documented history of their ancestors – while at the same time contrasting with the commonly-claimed history by many SpnCRL speakers themselves, and others – is demonstrated in Section 3.1.

3.1   The documented history of Saipan migration

Although the oral history of migration to Saipan credits the Satawalese navigator Agharubw with being the first Carolinian to seek Spanish permission to establish

---

16. The original German reads: *Aus dieser Verschiedenheit ihrer Herkunft erklärt sich auch der Reichtum an sinnverwandten Wörtern in der Sprache der Saipan-Karoliner, der diese wieder zum Verständigungsmittel geeigneter macht als einen der übrigen karolinischen Dialecte.*

a Saipan settlement, there is no historical documentary evidence linking the specific name Agharubw or Nguschul to the first colony. The writings of Arago (1823 PtI: 284, PtII: 12–13) do make it clear, however, that a navigator chief and his young family had settled on Saipan in 1815. There are also Spanish records stating that permission was both requested and given to a Carolinian delegation (Driver & Brunal-Perry 1996). There is no evidence that the delegation was from some place other than Satawal – nor that it was not Agharubw who led it. There is a specific atoll, however, that is identified in Arago's text – as interpreted by recent writers such as D'Arcy (2006: 159). Arago writes about a certain number of Carolinian survivors of an ill-fated voyage as being primarily from Lamotrek. They arrived on Saipan assumedly after Agharubw, although the particular month that Agharubw settled on Saipan is not recorded. And in 1818, as described by Hezel (1983: 106), a hundred Lamotrek islanders received permission to settle on Saipan. That addition of many Lamotrekese to the young settlement is further strengthened by the 1954 work of Spoehr. Among other information provided through scientific expeditions in the early 1800s Spoehr (1954: 70) writes, "… *Freycinet (1829–1837, vol. II: 327) stated that on Saipan in 1819 houses were just beginning to be built, four already being occupied by Carolinians from "Lamoursek" (probably Lamotrek) …"*.

Little information is available regarding migration events in the 1920s but there is no reason to not assume that canoes from most any western Caroline atoll would sometimes appear on Saipan during those early years. Spanish documents (Driver and Brunal-Perry 1996: 7) record a series of migrations forced by typhoons between 1838 and 1839 that added over a hundred migrants to Saipan. And again from 1847 to 1849 a series of typhoons resulted in many arrivals to Saipan from the outer islands D'Arcy (2006: 161). A super-typhoon in 1849 created a tsunami that completely covered the islands of Satawal and Lamotrek for several hours. Over forty survivors from those two islands eventually made their way to Saipan (Driver & Brunal-Perry 1996: 12–13).

By mid-century at least 300 outer-islanders were living on Saipan at any one time. There are indications that some of these settlers came from other islands in the Woleai region beyond Lamotrek. The main population though was made up of Satawalese and Lamotrekese, with the Lamotrekese migrants clearly dominating in number. This meant that the blended speech of the population was also dominated by the language of Lamotrek (that is, WOL)[17] … something that

---

17. The Lamotrekese interviewed during visits to Guam and Yap in 2009 tell me there are only very slight differences between their speech and that of those living on Woleai. Wordlists from the German South Pacific Expedition of 1908–1910 (Krämer 1937) reveal many orthographic differences between Lamotrekese and Woleaian but not a significant number of lexical differences.

would be expected since the social class of the Lamotrekese was higher than that of the Satawalese (see Hunter-Anderson & Zan 1996 for a description of the *sawei* which was a historical system of exchange and tribute that nurtured social ranking based on an atoll's respective geographical distance from the high-caste Yapese benefactors of the Gagil district).

The dynamics of population and speech in the Carolinian settlement on Saipan were changed considerably in the year 1865 and after when a British entrepreneur transported a large number of Carolinians to the Marianas. D'Arcy (2006: 162, quoting Farrell 1991) capsulizes those migration events between 1865 and 1869 this way:

> He [Johnson] initially hired 265 Carolinians from Pulusuk in 1865. Over the next four years he hired another 1,234 from Namonuito. … All the workers were hired from islands under threat of raids and tribute demands from Puluwat.[18]

It is difficult to trace exactly the destinations in the Marianas to which all the above-mentioned plantation workers from Pulusuk and Namonuito were sent. However, from a thorough inspection of the D'Arcy (2006), and particularly Driver & Brunal-Perry (1996: 137–138), and including Spoehr (1954: 71), the greatest likelihood is that the 265 Carolinians, and their subsequent offspring, eventually ended up in the Arabwal (Garapan) settlement sometime after 1889 – twenty-four years after the 1865 displacement of the Pulusukese to various islands in the Marianas. As to the large population from Namonuito, it is quite clear from other sources (including Alkire 1984: 272) that those members of the migrant-workers population (who did not return to Namonuito) were the ones responsible for establishing the new settlement, Tanapag, a couple miles north of Garapan. Alkire states clearly that, "In 1870, the Carolinian community was further augmented, when 230 people from Namonuito arrived by way of Tinian and established themselves in the village of Tanapag, north of Garapan."

The language that developed from the Namonuito settlers is today called Northern Saipan Carolinian, or, as more preferred, Talaabog (TAL), as explained earlier in Note 6. The intelligibility figures displayed in Table 1 show that the remaining fluent Talaabog speakers alive today – of which there are very few – speak a language that can be considered a close dialect of the Namonuito language. The fact that Talaabog was created from migrants from the one language – Namonuito – means that Talaabog is not a language blended from other Carolinian languages. Only the Southern Saipan Carolinian language is here considered a blended language.

---

18. Current spelling of this atoll is Polowat.

The conclusion derived from my research (Ellis 2012: 378) on the migration of Carolinians to Saipan is that those islanders apart from the Satawalese islanders arriving before 1850 were primarily from Lamotrek and other islands between Lamotrek and Woleai – all of whom spoke WOL. Whereas migrants arriving after 1850 were primarily from Pulusuk and islands directly north thereof (see Figure 4) – all of whom spoke POL.

3.2 The parallel evidence of the Saipan Carolinians migration recorded in historical records and the record preserved within the linguistic code of today's SpnCRL speakers

As demonstrated in Section 2, the SpnCRL language – which finally started to develop a standard form during the first decade of the 1900s – is a blended language from WOL, SAT, and POL; with the phoneme inventory being influenced most by WOL and the lexicon being influenced most by POL. This section examines the likely sociolinguistic reasons for SpnCRL's particular division of linguistic affiliation between WOL and POL.

It is clear from evidence given in Section 3.1 that the language setting of the earlier wave of migration was dominated by the language of migrants from Lamotrek – speakers of WOL. It is also clear that later waves of migrants were dominated by Pulusuk speakers – speakers of POL. Putting this in context of the intelligibility study, the results show that SpnCRL speakers have a much higher level of intelligibility with POL than they do with WOL. It is not surprising in light of those results that the lexicon of SpnCRL would be also much more similar to POL than it is WOL. The question is, in view of this, why have the marked phonemes of WOL managed to survive all through the years in spite of the extensive POL influence on SpnCRL?

Already mentioned is that Woleaians and Lamotrekese have a higher social status than Polowatese. It makes sense, then, that at the time SpnCRL finally emerged as a new language early in the 1900s the Saipan Carolinians would have preferred to pronounce words in the way that WOL speakers do than in the way that POL speakers do. But there is also a second factor that could have had a very significant effect on language use during the century of Saipan migration. The term "Founder Principle," is used by Mufwene (1996) (also a contributor to this volume) to explain the phenomenon where the initial language community in a newly-established society has a long-term and disproportionate impact on the language ecology over that of subsequent language communities that also become part of the society. The concept of the Founder Principle, used more broadly in

respect to a socio-cultural ecology, is described some years earlier by Zelinsky (1973: 13–14) in a way that applies well to the settlement situation on Saipan.

> Whenever an empty territory undergoes settlement … the specific characteristics of the first group able to affect a viable self-perpetuating society are of crucial significance to the later social and cultural geography of the area, no matter how tiny the initial band of settlers may have been …

Although SAT speakers were most likely the first group of people starting to settle permanently on Saipan it was the WOL speakers of Lamotrek that were the *"first group able to affect a viable self-perpetuating society."* The first houses built, for instance, were apparently inhabited by Lamotrekese, as earlier noted. It is also true that the much later migrations of POL speakers succeeded in dominating the SpnCRL's lexicon. Nonetheless, it is the WOL pronunciation, even for a large number of POL-contributed lexical items, that has been largely retained by SpnCRL speakers.

While phonemes are more stable than lexical items in general, the order of migration may further explain why the marked phonemes of WOL have managed to survive through the years in spite of the fact that both the levels of intelligibility and the percent of lexical similarity favor POL.

What does this imply? It implies that the historical documentation of the Saipan migrations is corroborated by the linguistic features of spoken SpnCRL. Such corroboration demonstrates that the commonly-believed migration history that persists on Saipan (and also the history suggested by earlier historical linguistics) does not actually correspond with the linguistic-historical documentation of the Saipan migrations. The significance of this is that were there to be no surviving historical documents covering the Saipan migration, the general features of that migration would nonetheless have been preserved on the tongue of the remaining speakers of traditional Saipan Carolinian. That alone would strengthen the argument that the benefits of thorough documentation of languages like SpnCRL, even with only a shallow human history, are well worth the costs.

4. Conclusion

Like many other languages that have emerged in recent times as fusions between two or more other languages or dialects, Saipan Carolinian is on track to have a very short life-span – perhaps just over a hundred years in total, for today it is difficult to find speakers below the age of thirty or so on Saipan who are still speaking Saipan Carolinian. Not surprisingly, what is heard mostly is English. And the elder traditional speakers of Saipan Carolinian are too-quickly disappearing: a

common story, unfortunately. There is nothing common, however, about the linguistic composition of the language, as this study has attempted to reveal.

In this paper the language of the Saipan Carolinians was introduced and it was demonstrated that their language eventually emerged as a standardized language around the first decade of the 1900s on the island of Saipan – an island uninhabited at the time of the original Carolinian settlement in the early 1800s. The language was also demonstrated to be a product of the other languages of the Carolinian continuum, located in the Caroline Islands of Micronesia – particularly three source languages, Woleaian, Satawalese, and Polowat. Particularly notable was how the speakers of this blended language managed at the end of the migration to end up with phonemes that were most like Woleaian and a lexicon that was most like Polowatese – two languages that are mutually unintelligible. Also presented in this study is the way in which the Saipan Carolinian lexicon is an *amalgamation* of the other languages' lexicons, thus building a far richer resource of synonyms than the other languages. The paper then touched on the oral history of the speakers of Saipan Carolinian and their contention that their ancestral migrants were primarily speakers of Satawalese; not the languages of the other atolls. That belief – on the part of the speakers of Saipan Carolinian, as well as previous linguists – that Saipan Carolinian is most closely affiliated with Satawalese, is at odds not only with the linguistic composition of Saipan Carolinian speech but also at odds with the many historical records that document how the majority of migrants came not from Satawal, but from the atolls in both the Woleai region and the Polowat region. In essence, then, the correct history of Saipan migration is not in the memory but rather on the tongue of SpnCRL speakers.

This paper supports the position that diverse sociolinguistic ecologies emerging from languages and dialects in contact have contributed greatly to the world community in expanding our understanding of how human language works. It is hoped the current and on-going study of Saipan Carolinian, in spite of its short history and its apparently short future, can also contribute something to our understanding of how language works.

References

Alkire, W.H. 1984. The Carolinians of Saipan and the Commonwealth of the Northern Mariana Islands. *Pacific Affairs* 57(2): 270–283. doi:10.2307/2759128

Arago, J. 1823. *Narrative of a Voyage Round the World in the Uraine and Physicienne corvettes commanded by Captain Freycinet, during the years 1817, 1818, 1819 & 1820*. London: Treuttel & Co.

Auer, P. 1998. From code-switching via language mixing to fused lects: Toward a dynamic typology of bilingual speech. *Interaction and Linguistic Structures Freiburg*, 6. <http://kops.uni-konstanz.de/bitstream/handle/123456789/3677/470_1.pdf>

Bakker, P. 1997. *A Language of Our Own: The Genesis of Michif, the Mixed Cree French Language of the Canadian Metis*. Oxford: OUP.

Casad, E.H. 1974. *Dialect Intelligibility Testing* [Summer Institute of Linguistics Publications in Linguistics and Related Fields 38]. Norman OK: Summer Institute of Linguistics of the University of Oklahoma.

D'Arcy, P. 2006. *The People of the Sea: Environment, Identity and History in Oceania*. Honolulu HI: University of Hawai'i Press.

Driver, M. & Brunal-Perry, O. (eds). 1996. *Carolinians in the Mariana Islands in the 1800s: Selected Documents from the Holdings of the Spanish Documents Collection at the Micronesian Area Research Center*. Division of Historic Preservation, Department of Community and Cultural Affairs, Commonwealth of the Northern Mariana Islands.

Ellis, S.J. 2007. Language bending. *Working Papers of the Department of Linguistics*. Mānoa HI: Department of Linguistics, University of Hawai'i.

Ellis, S.J. 2012. Saipan Carolinian, one Chuukic Language Blended from Many. PhD dissertation, University of Hawai'i at Mānoa.

Ellis, S.J. & Fruit, C., in collaboration with J. Limes & D. Omar. 2003–current. *Saipan Carolinian–English Dictionary in Toolbox* (digital file). Saipan: Carolinian Affairs Office, CNMI.

Farrell, D.A. 1991. *History of the Northern Mariana Islands*. Saipan: Public School System of the Northern Mariana Islands.

de Freycinet, L.C.D. 1829. *Voyage Autour du Monde...: Exécuté sur les Corvettes de S.M. L'Uranie et la Physicienne, pendant les Années 1817, 1818, 1819 et 1820*. Paris: Pillet Aâné.

Fritz, G. 1911. *Die Zentralkarolinische Sprache; Grammatik, Übungen u. Wörterbuch der Mundart der westlich von Truk liegenden Atolle, insbesondere der Saipan-Karoliner*. Berlin: Georg Reimer. doi:10.1515/9783111728582

Grimes, J.E. 1992. Correlations between vocabulary similarity and intelligibility. In *Windows on Bilingualism*, E.H. Casad (ed.). Dallas TX: University of Texas at Arlington and SIL.

Hezel, F.X. 1983. *The First Taint of Civilization: A History of the Caroline and Marshall Islands in Pre-colonial Days, 1521–1885*. Honolulu HI: The University of Hawai'i Press.

Hoffman, C.W. & USMC. 1950[1965]. Saipan: The beginning of the end. In *A Chronology of the United States Marine Corps 1935–1946*, Vol. II: U.S. Marine Corps, C.A. Tyson (ed.). Washington DC: History and Museums Division Headquarters.

Hunter-Anderson, R.L. & Zan, Y.G. 1996. *Demystifying the Sawei, a traditional interisland exchange system*. ISLA: A journal of Micronesian Studies 4. University of Guam: University of Guam Press.

Jackson, F. 1983. The Internal and External Relationships of the Trukic Languages of Micronesia. PhD dissertation, University of Hawai'i at Mānoa.

Jackson, F. & Marck, J., in collaboration with J. M. Elameto. 1991. *Carolinian-English Dictionary*. Honolulu HI: University of Hawaii Press

Krämer, A. 1937. Zentralkarolinen. 1. Halbband: Lámotrek-Gruppe. In *Ergebnisse der Südsee-Expedition 1908–1910, II. Ethnographie: B. Mikronesien* 10, G.T. Oleai-Feis (ed.). Hamburg: Friedrichsen, De Gruyter.

Lewis, M.P. (ed.). 2009. *Ethnologue: Languages of the World*. Dallas TX: SIL International.

Lewis, M.P., Simons, G.F. & Fennig, C.D. (eds). 2014. *Ethnologue: Languages of the World*. Dallas TX: SIL International.

Living Tongues Institute for Endangered Languages. 2007 to present. Mission Statement. 16 October 2014. <http://www.livingtongues.org/>

Marck, J.C. 1986. Micronesian dialects and the overnight voyage. *Journal of the Polynesian Society* 95(2): 253–258.

McCoy, M. 1976. A renaissance in Carolinian-Marianas voyaging. *Polynesian Society Memoir* 39. Wellington.

Mufwene, S.S. 1996. The founder principle in creole genesis. *Diachronica* 13(1): 83–134. doi:10.1075/dia.13.1.05muf

Odango, L.E. 2015. Afféú fangani 'join together': A Morphophonemic Analysis of Possessive Suffix Paradigms and a Discourse-based Ethnography of the Elicitation Session in Pakin Lukunosh Mortlockese. PhD dissertation, University of Hawai'i at Mānoa.

Quackenbush, E.M. 1968. From Sonsorol to Truk: A Dialect Chain. PhD dissertation, University of Michigan.

Spoehr, A. 1954. *The Ethnology of a War-Devastated Island*. Chicago IL: Fieldiana: Anthropology, Chicago Natural History Museum. doi:10.5962/bhl.title.5465

US Census Bureau. 2002. Population and Housing Profile: 2000. CNMI. US Department of Commerce. https://www.census.gov/prod/cen2000/island/CNMIprofile.pdf.

Trudgill, P. 1986. *Dialects in Contact*. Malden MA: Blackwell.

Weinreich, U. 1953. Languages in contact: Findings and problems. *Publications of the Linguistic Circle of New York* 1.

Wikimedia Commons. 27 October 2010. Pacific Cultural Areas (map). 30 October 2014. <http://upload.wikimedia.org/wikipedia/commons/9/93/Pacific_Culture_Areas.jpg>

Zelinsky, W. [1973]1992. *The Cultural Geography of the United States: A Revised Edition*. Englewood Cliffs NJ: Prentice Hall.

Aikanã and Kwaza

Their ethno-historical and sociolinguistic context in Rondônia, Brazil

Hein van der Voort

Museu Paraense Emílio Goeldi, Belém

Aikanã and Kwaza are highly endangered language isolates in southeastern Rondônia, Brazil. Today, their speakers live in indigenous reserves with the last speakers of Latundê (Northern Nambikwaran) and Salamãi (Mondé, Tupian). Whereas the elderly try to maintain indigenous cultures and languages, younger generations are increasingly oriented towards Western lifestyle, and Portuguese is becoming the family language. Protection of indigenous heritage is guaranteed in the Brazilian constitution, but the authorities have little knowledge about Indians and languages, and indigenous communities are sometimes indifferent or even divided over the issue. Aikanã will probably survive this century. The future of Kwaza is unclear. Documentation and description of languages and cultures can be adapted to the community's needs and play a role in preservation.

Keywords: Aikanã, Amazonian languages, Brazil, documentation, isolates, Portuguese, preservation, Rondônia

1. Introduction

This article concerns the indigenous people of southeastern Rondônia, Brazil. It focuses on the Aikanã and the Kwaza, their history, culture, language, and their relationships with neighbouring peoples and with their ecological context. Both the Aikanã language and its immediate neighbour, Kwaza are genetic isolates that are highly endangered (175 and 25 speakers, respectively). The region where they are spoken is under high cultural and environmental pressure from the national society. The (pre-)histories of the Aikanã and Kwaza speakers are considerably intertwined. The languages show much evidence of contact, although an ancient genetic connection cannot be excluded.

The present article reports on ongoing efforts to create an extensive archive of transcribed, translated and annotated audio and video recordings of linguistic, cultural, ethnobiological, archeological and historical importance concerning the Aikanã people and their language. In addition, the existing archive of Kwaza documentation, which resulted from one of my earlier projects is considered, as it is being incorporated and complemented by the same ongoing modern systematic documentation efforts as for Aikanã. Finally, an ethnohistorical study of the Aikanã and Kwaza peoples based on field recordings and existing documents, similar to the recent work of Reesink (2012) on the neighbouring Nambikwara groups, is a desideratum.

These efforts have the potential to fill the ethnohistorical gap between the Nambikwara context to the east (see Lévi-Strauss 1955; Reesink 2012), the Tupari and Jabuti context to the west (see Caspar 1975; Snethlage 1937; Brijnen & Adelaar 2010; van der Voort 2008a,b), the Mekens and Kanoê context to the southwest (see Galucio 2006; Bacelar 2003) and the Mondé context to the northeast (see Mindlin 1985; Moore 1984). Linguistic aspects can be linked with existing comparative lexical and grammatical feature databases (e.g., Muysken et al. 2014), which are instrumental in the reconstruction of the history of language contact and intercultural communication in the Guaporé-Mamoré linguistic area (see Crevels & van der Voort 2008). The envisioned outcome of these efforts can form the basis for further studies evaluating a possible genetic relationship between Aikanã and Kwaza (see van der Voort 2005).

The ongoing project is funded by the Volkswagen Foundation within the DoBeS (Dokumentation Bedrohter Sprachen) programme. It is an interdisciplinary project that presently involves the anthropologist Lisa Katharina Grund, the indigenous research assistant Cândida Aikanã, and the linguists Joshua Birchall and myself. It provides an opportunity to fulfil the wishes of the Aikanã and Kwaza communities to codify their unique languages and preserve their use in modern times. That being also their right according to the Brazilian constitution, the preliminary results of the present project have already proven to be of immediate relevance as an instrument for both governmental and non-governmental initiatives to provide and support diversified education for the communities. In addition, the project has been aiding the communities in preserving their cultural and ecological knowledge and memory. In short, the project and its growing digital archive have been benefitting the indigenous communities in their efforts to preserve and revitalise their language and culture, as well as the scholarly community with regard to understanding the cultural and linguistic relationships in this fascinatingly diverse region.

2.   General background

The Brazilian federal state of Rondônia, about the size of the U.K., is located in the southwestern reaches of the Amazon basin, on the border with Bolivia. South-eastern Rondônia is one of the most degraded parts of the Amazon region today. Apart from certain original areas of savannah on the Parecis plains, Rondônia was almost completely covered with rainforest vegetation until the 1960s. From then on, the building of the BR-364 highway from the south to the north in com-bination with governmental land distribution policies have led to uncontrolled deforestation and colonisation by representatives of Western culture, which have accelerated the physical and cultural extinction of the aboriginal inhabitants. To-gether with parts of the state of Pará, Rondônia represents a prime example of large-scale irreversible ecological destruction in the Amazon. During the past 50 years, it has lost more than half of its natural vegetation. The indigenous popula-tions have dwindled in numbers due to exogenous diseases, violence and forced resettlement. Their remnants are now mainly located in ethnically diverse indig-enous reserves, sometimes far away from their traditional homelands, and many are now also trying to make a Western-style living in the slums of local boom towns. The Portuguese language is the lingua franca almost everywhere and is gradually also acquiring the function of a first language. Even though there is more public ecological awareness than there used to be, policies continue to be defined by notions like 'progress' and 'development' in supposedly unused territo-ry. Nowadays, so-called development can even be seen from space. For example, through Google Earth (coordinates 9°–13° S with 60°–64° W) one can see the north-south network of roads, fields and towns progressively expand towards the west and the east. The southeast of Rondônia is especially affected. On the official maps of Rondônia, natural parks and indigenous reserves are outlined and indi-cated by a contrasting colour. Today, their almost exact shapes can also be seen on satellite photos, forming green patches in a sea of yellow and grey (in spite of federal environmental laws that prohibit deforestation within 10 kilometres of the demarcation lines that define parks and reserves).

Rondônia is one of the most linguistically diverse parts of South America. The state is home to approximately 25 languages, divided into five linguistic fami-lies and three language isolates. Moreover, it can be regarded as part of a linguistic area that includes sections of Mato Grosso and the Bolivian Amazonian lowlands (Crevels & van der Voort 2008: 154). One can assume that this linguistic diversity more or less reflects the original situation, albeit extremely threatened by extinc-tion. Because the physical and cultural changes sketched above have taken effect relatively recently, there are still elderly persons who were born and raised in the

traditional context.[1] They tend to speak the indigenous languages, have memories of the traditional material, intellectual and spiritual cultures, and may still try to pass aspects of their knowledge on to the younger generations. One could estimate that in traditional times, an indigenous ethnolinguistic group consisted of perhaps between 1,000 and 5,000 individuals. Nowadays more than half of the indigenous languages have less than 50 speakers, and one-third of the languages have fewer than ten speakers. Table 1 lists the indigenous languages of Rondônia, with their genetic classification, population numbers, speaker numbers and references to major linguistic work.

Table 1. Linguistic diversity of Rondônia[2]

	Language	Family	Population	Speakers	Major recent work
1	Kaw Ta Yo (Kuyubi)	CHAPACURA	55	2	Duran 2000
2	Miguelenho		50	1	
3	Oro Towati (Oro Win)		56	5	França 2002
4	Wari' (Pakaanova)		2700	2700	Everett & Kern 1997
5	Arikapu	MACRO-JÊ	30	1	van der Voort 2010
6	Djeoromitxi	(*Jabuti*)	165	30	Pires 1992; Castro 2012
7a	Latundê	NAMBIKWARA	19	19	Telles 2002
7b	Lakondê	(*Northern*)	7	1	
8	Kaxarari (Kaxariri)	PANO	322	300	
9	Uru-eu-wau-wau (Amondawa, Jupa'u)	TUPI (*Tupi-Guarani, Kawahib*)	183	183	Sampaio 2001
10	Karipuna		14	10	
11	Karitiana	(*Arikém*)	320	320	Storto 1999; Everett 2006
12	Purubora	(*Puruborá*)	62	2	Galucio 2005; Monserrat 2005

---

1. I use the term 'traditional' in the sense 'before intensive contact with Western society'. It is likely that there have been effects of (indirect) contact with Westerners long before the elderly community members were born, possibly even as early as the 16th century.

2. Population numbers were estimated on the basis of censuses by the Instituto Socioambiental (ISA) from 2009 and the national health foundation FUNASA from 2006 and by personal communication with colleagues in the field. See also Crevels (2012). Classification in between brackets in italics concerns subfamilies.

Table 1. (*continued*)

	Language	Family	Population	Speakers	Major recent work
13	Tupari	(*Tupari*)	433	150	Alves 2004; Rodrigues & Caspar 1958
14	Makurap		381	50	Braga 2005
15	Mekens (Sakirap)		84	22	Galucio 2001
16	Wayuru (Ajuru)		94	8	Nogueira 2011
17	Akuntsu (Akũtsũ)		5	5	Aragon 2008
18a	Aruá	(*Mondé*)	36	12	
18b	Cinta Larga		645	645	
18c	Gavião		523	523	Moore 1984
18d	Zoró		599	599	
19	Salamãi (Mondé)		10?	2	
20	Surui-Paitér		1007	1007	van der Meer 1982; Guerra 2004
21	Karo (Arara)	(*Ramarama*)	208	200	Gabas Jr. 1999
22	Aikanã (Masaká, Kasupá, Huari)	ISOLATE	200	175	Vasconcelos 2002
23	Kanoê (Kapixaná)	ISOLATE	95	3	Bacelar 2004
24	Kwaza (Koaiá)	ISOLATE	40	25	van der Voort 2004
25	Isolado do Tanaru	UNCLASSIFIED	1	1	
26	Isolados do Massaco	UNCLASSIFIED	100	100	
27	Akuntsu-Kanoê	PIDGIN	–	7	

Anthropologists have identified several cultural complexes in Rondônia. Lévi-Strauss (1948: 371) identified the cultures on the right side (going downriver) of the Guaporé River as belonging to the 'Guaporé culture area,' divided into a western Chapacura sub-area and an eastern Tupi sub-area. This classification was superseded by the research of Denise Maldi, who has conducted ethnohistorical research in western Rondônia and Mato Grosso since the early 1980s. Maldi (1991) defines a part of Lévi-Strauss' Tupi area as the 'Marico cultural complex of Rondonia.' The Marico cultural complex includes cultures of Tupi-speaking peoples, Jabutian-speaking peoples, and peoples speaking isolate languages, and share the following traits:

– semi-nomadic swidden agriculture combined with hunting and gathering
– relatively small egalitarian societies

- territorial subgroups that often bear animal names
- territorial subgroups that may form alliances despite linguistic differences
- spiritual culture involving shamanism and hallucinogenic substances
- material culture characterized by the *marico*, a crochet bag made of the fibers of specific palm tree leaves
- consumption of *chicha*, a fermented alcoholic brew mainly based on maize, yam, manioc, *patauá* seeds (of the *Oenocarpus bataua* palm tree) or banana, mashed, fermented and strained in a specific way

In spite of the great linguistic diversity, these shared cultural traits point to a long history of interethnic contacts and intermarriage between neighbouring groups. In addition to cultural correspondences, the diverse languages often show lexical borrowings characteristic of the region and areally-diffused grammatical traits (Crevels & van der Voort 2008: 166ff.). Interethnic contact was not limited to specific river basins but also took place over land between groups on different headwaters. Although contributing to different river basins, the headwaters of the Branco, Mekens, Tanaru and São Pedro rivers all start relatively close to each other on the same section of the Parecis plains. The languages spoken by groups belonging to the Marico cultural complex are listed in Table 1 with numbers 5, 6, 12–18a, 19, 22–24.

For some ethnic groups, especially those living in proximity to the first Western settlements, cultural and demographic changes have been more invasive and thorough than for other groups in remoter areas, where Western influence has been encroaching only recently. The southeast of Rondônia was one of the first regions to enter into contact with representatives of Western culture. The first contacts probably took place in the 18th century, when Portuguese explorers sailed up the Madeira River and established military forts on the Guaporé River to defend the empire against Spanish invasion (see Maldi 1989).[3] During the same epoch, gold prospectors came over land from the direction of Cuiabá in the east, and their brief presence has left its marks in the form of fabled rumours about the goldmines of Urucumacuan (see Dequech 1943).[4] The interethnic contacts of this period were rarely documented and only a few reports of skirmishes

---

3. The ruins of the Forte Príncipe da Beira near the smuggler's town of Costa Marques, close to the mouth of the Cautário River, were preserved and can be seen on Google Earth (12°25′40″ S. & 64°25′20″ W.).

4. Throughout the 20th century, mineralogists (e.g. Dequech) have searched for these legendary mines in vain until they were possibly found a few years ago in the Roosevelt River basin. This has led to a frenzied diamond rush that has brought violence and disease upon the reserve of the Cinta Larga and has cost the lives of many Indians and Westerners.

with local Indians and lists of unidentifiable names of possible ethnicities have been preserved in the archives (see Price 1983). The first permanent contacts with Westerners in southeastern Rondônia were established during the rubber boom around the turn of the 19th century. Rubber entrepreneurs and exploiters of other important forest resources, such as the root of *ipecacuanha* (once a popular emetic to induce vomiting in case of poisoning), employed and sometimes enslaved Indians of the region up until the 1960s. During this period several ethnic groups were almost wiped out by diseases such as influenza and measles, against which they had no immunity. When Snethlage (1937: 161ff.) visited the Arikapu in the 1930s they were already a devastated nation. Also the Kwaza, Kanoê, Salamãi, Wayoró and Aruá were decimated during these years. The elderly indigenous persons of the region are the survivors of this mass extermination, and they all can give gruesome eyewitness reports.

After the 1950s, when the Brazilian indigenous population numbers were at an all-time low on a national level (see Gomes 1988), most indigenous populations of southern Rondônia were resettled in indigenous reserves, officially protected by the National Foundation for the Indian (FUNAI),[5] and administratively integrated into mainstream society. The enlightened constitution of 1988 guarantees indigenous populations the right to inhabit their original lands and to preserve their cultures and languages. However, there is a big gap between theory and practice, especially in the remoter parts of Brazil, such as Rondônia. Illegal lumbering, mining and religious proselytisation, sometimes with the consent or even the participation of local authorities, form a constant pressure on indigenous lands and cultures. There are still uncontacted groups in Rondônia, albeit very small groups. Their unmarked territories are invaded illegally by loggers, farmers and mineral prospectors, and they risk being killed by diseases and violence. Those that have figured in the international press, such as the Omeré Indians (two small groups of Kanoê and Akuntsu first contacted in 1995) and the lone "Isolado do Tanarú" (the unidentified Indian on the headwaters of the Tanaru River, who avoids all contact; see Holtwijk 2006; Reel 2010), are somewhat protected by their fame, but the responsible FUNAI agents have to wage a permanent battle with the local authorities and the surrounding Western population for the rights of these extremely vulnerable groups.

Map 1 shows Rondônia and the indigenous reserves where some of the languages mentioned in this article are spoken.

---

5. Fundação Nacional do Índio (formerly S.P.I., Serviço de Proteção aos Índios).
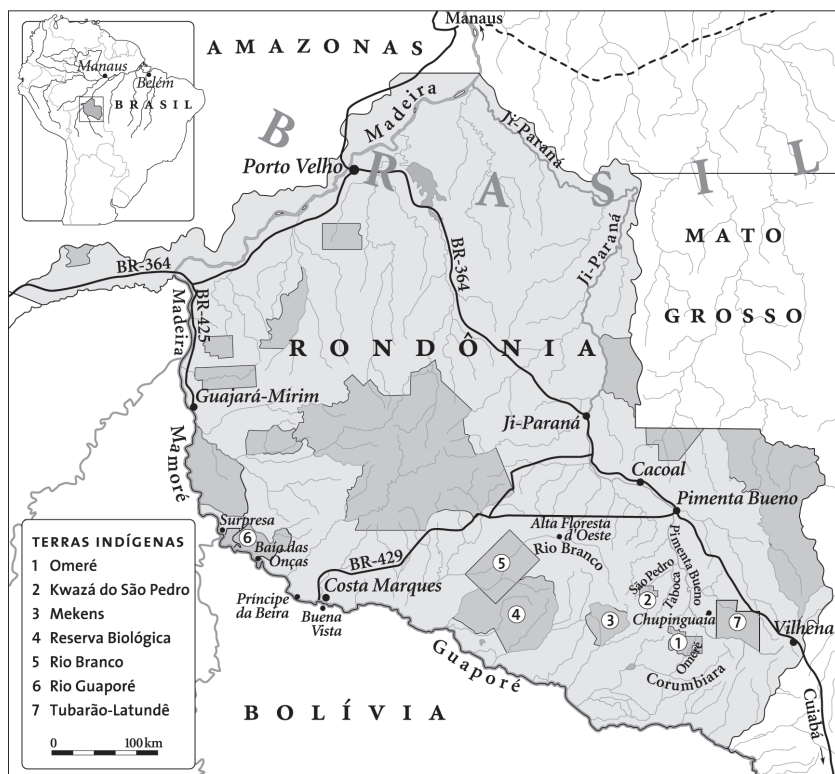
Map 1. Rondônia and its indigenous reserves (from van der Voort 2004)

In the numbered reserves on the map the following indigenous languages are spoken nowadays (the numbers in parentheses that follow the language names correspond with the numbers of Table 1):

1 – Akuntsu (17), Kanoê (23)
2 – Aikanã (22), Kwaza (24)
3 – Mekens (15)
4 – unknown (26)
5 – Djeoromitxi (6), Tupari (13), Makuráp (14), Aruá (18a)
6 – Arikapu (5), Djeoromitxi (6), Tupari (13), Makuráp (14), Wayoro (16), Kanoê (23)
7 – Latundê (7a), Salamãi (19), Aikanã (22), Kwaza (24)

The Aikanã and Kwaza languages, which represent the central focus of the documentation activities described in the present article, are spoken in reserves number 2 and 7.

3.   History of ethnolinguistic research in southeastern Rondônia

Whereas indigenous peoples on the Bolivian side of the Guaporé had been in permanent contact with Westerners, and Spanish Jesuits began to document their languages at the end of the 17th century, knowledge of the peoples on the Brazilian side was practically nil until the end of the 19th century. One of the first to document indigenous languages and peoples of southeastern Rondônia was the army captain and medical surgeon João Severiano da Fonseca. In 1877 he sailed down the Guaporé River and noted down word lists from local groups, including the now extinct Palmellas, who spoke a Cariban language (Fonseca 1881:190ff.).

The first systematic ethnographic observations of Nambikwara and various other groups of western Mato Grosso and eastern Rondônia, as well as language samples, were registered by members of the Comissão Rondon (1907–1915) during the building of a telegraph line from Cuiabá to Porto Velho. On separate expeditions in 1914 the explorer Percy Fawcett and the ethnographer Erland Nordenskiöld met ethnic groups on tributary rivers of the right side of the Guaporé and documented aspects of the indigenous culture and language, including Arikapu by Fawcett (1915; Rivet 1953; van der Voort 2012) and Aikanã by Nordenskiöld (1915:363–398). About twenty years later, in 1933–1935, the German ethnographer Emil Heinrich Snethlage (1937) explored various tributaries on both sides of the Guaporé River and documented many indigenous languages and cultures, including Tupari, Wayoró, Mekens, Aruá, Djeoromitxi and Arikapu. In 1937 the anthropologist Claude Lévi-Strauss (1955) worked for a year among Nambikwara groups. He returned to São Paulo by descending the Pimenta Bueno River in 1938 with the Brazilian ethnographer Luiz de Castro Faria and documented various languages and cultures on the way, including Salamãi (see Castro Faria 2001) and Kwaza (van der Voort 2004:30–33, 2008c). A few years later, in 1941–1943, expeditions of the federal Indian Protection Service S.P.I. (see Zack 1943) and the National Department of Mineral Production[6] (see Dequech 1943) traversed southeastern Rondônia and contacted and worked with its indigenous populations. In 1948, the Swiss ethnographer Franz Caspar (1953) paid a brief visit to the Rio Branco region and returned in 1954–1955 to live for a year with the Tupari (see Caspar 1975). During his second visit Caspar was one of the last to document indigenous groups of southeastern Rondônia more or less in their traditional cultural context and original habitat.

In recent years, there have been encounters with isolated indigenous groups in this region, such as the abovementioned Omeré and Tanaru Indians, but only

---

6.   Both governmental organisations were at the time part of the Ministry of Agriculture.

as refugees on small islands of rainforest in the midst of endless pastures and soybean fields. There is one larger isolated group of about 100 members in the Massaco protected area, who defend their territory with lethal traps and avoid any form of contact. Nothing is known about their language, and the only basis for their attempted identification as 'Siriono' or 'Tupi-Guarani' are their long bows. Nowadays, standard FUNAI policy is not to seek contact unless the Indians are threatened by external factors.

Although the remainders of the indigenous peoples of the region have now mostly been resettled in multi-ethnic federal reserves, where traditional aspects of their cultures are progressively disappearing, and their languages are becoming extinct, the last decades have seen an increase in linguistic studies. As shown in Table 1 above, many languages have recently received attention, which is leading to substantial theses, articles and in some cases published language descriptions. Many initiatives focus on Tupian languages, although the smaller families and isolates are now also being covered. The Nambikwaran language family is the subject of a research programme organised by Leo Wetzels of the Vrije Universiteit in Amsterdam. The Chapacuran languages are studied among others by the linguist Joshua Birchall at the Museu Goeldi in Belém, and my own work on the Jabuti languages is ongoing. The isolate language Kanoê was described and is now being documented by Laércio Bacelar as part of the UNESCO-administered documentation initiative of the Museu do Índio (FUNAI). These initiatives are all just in time, before the complete disappearance of most of these languages. Of the gaps in language documentation that remain, I regard those concerning Aikanã and Kwaza as extremely urgent.

4.   The modern history and present situation of the Aikanã

The Aikanã language is spoken by around 175 individuals. The majority of its speakers live in the indigenous reserve Terra Indígena Tubarão-Latundê. All generations learn the language, but traditional cultural knowledge is disappearing under Western ecological, cultural and religious pressure. For this reason, the elderly people tend to be the best informants, some of whom were born in the indigenous communal house, before the Second World War.

The Aikanã lived originally on the headwaters of, and along, the Pimenta Bueno (or Apediá) River in southeastern Rondônia. Their neighbours in those times were the Kwaza (or *Koaiá*), Kanoê, Salamãi and Kepkiriwat (an extinct TU-PIAN language), and, at a somewhat greater distance, the Latundê and the Mekens (or *Sakurabiat*).

The Aikanã can be encountered in the literature by different names: *Aikanã*, *Massaká*, *Kassupá*, *Huari*, *Corumbiara*, *Mondé*, *Tubarão*, and there are still other names. However, the autodenomination is *Aikanã*. The Aikanã were mentioned for the first time in 1913, by the name *Uapuruta* (recognised by the Aikanã as *Waikurutá*), on a hand-drawn map based on a visit to the Kepkiriwat Indians by the Comissão Rondon (cf. Rondon & Faria 1948: 183). Aikanã informants today confirm that the *Waikurutá* represented one of the Aikanã subgroups. Apparently, Rondon considered them, like the *Charamein* (Salamãi), to be a subgroup of the Kepkiriwat, probably because they maintained friendly relationships with each other and because their cultures were similar (Anonymous 1916: 334).

The first clear eyewitness account of the Aikanã people, with descriptions, photographs and a sample of their language, was given in 1914, by the Swedish ethnographer baron Erland Nordenskiöld. In his 1915 book Nordenskiöld describes his encounter with the "Huari" people on the headwaters of the Corumbiara River. The ethnonym *Huari* was given to them by the (now extinct) Pauserna people. The personal names, as well as the 83 "Huari" words that were registered in Nordenskiöld's manuscripts[7] are practically identical with those of the present-day Aikanã.

The Aikanã had probably been in contact with Westerners already before the short visit of Nordenskiöld, but the big cultural changes only came a little later, with the entry of the rubber entrepreneur Américo Casara (see Albert 1964), who employed them in the extraction of rubber, ipecacuanha and other native products of the region. There are several indications that the Aikanã got along reasonably well with Westerners. They often provided valuable services to the expeditions of the Comissão Rondon in the first decade of the 20th century, the Urucumacuan Expedition of the 1940s, the building of the BR-364 highway in the 1960s, and other government initiatives until present times. Currently they are participating in FUNAI expeditions to protect isolated groups in the Massaco, Omeré, Tanaru and other regions.

Since 1973, the majority of the Aikanã have been living on the reserve Terra Indígena Tubarão-Latundê (demarcated in 1983), close to the town of Chupinguaia. There they continued to extract rubber until 1997 and subsisted by hunting and swidden agriculture. In 1975 they contacted an isolated family of Aikanã, led by 'Capitão'[8] Arui Uhunei, who had split off in the late 1940s and happened to live in a remote part of the reserve. With the re-established relationships with

---

7.   Acessible in the archives of Världskulturmuseet in Gothenburg. The word list published in Nordenskiöld (1915: 371–372) represents a selection of 67 entries.

8.   The honorific expression *Capitão* 'captain' stems from the early contact period, when Brazilian officials used this term to refer to leading figures in the communities they encountered.

this family, knowledge of indigenous shamanic healing was reintroduced into the community, as well as new marriage candidates. In 1977, the Aikanã contacted a previously unknown group of Northern Nambikwara, the Latundê, in the remote savannah-like eastern interior of the reserve. Nowadays the Aikanã live on this reserve together with the remaining representatives of the Kwaza people, one semi-speaker of Salamãi and with the last of the Latundê.[9] Today, speakers of Aikanã also live in the Terra Indígena Kwazá do Rio São Pedro, in the nearby towns like Chupinguaia and Vilhena, and in the capital Porto Velho.

One of the negative and perverse aspects of the (direct or indirect) contact with Westerners was the decimation of the Aikanã, especially through contagious diseases against which they had no resistance. Furthermore, the loss of their most fertile lands and cultural pressure from representatives of Western culture (lumbermen, missionaries, government officials, etc.) has resulted in a decrease in the transmission of aspects of indigenous culture.

5.   Previous research and documentation of Aikanã language and culture

Following Nordenskiöld's first record, various interested laymen (especially S.P.I. officials) registered samples of the languages of the region. In particular, the report (and its appendix) of Estanislau Zack (1943), leader of the aforementioned S.P.I. expedition, deserves mention here. In the 1950s several foreign ethnographers documented parts of the language and oral traditions of the Aikanã; among these were Etta Becker-Donner (1955) and Wanda Hanke (1956). In the 1960s the SIL missionary linguists Willem Bontkes (1968) and Wilbur Pickering (1968) registered small vocabularies of the language. When preparing the demarcation of the Aikanã indigenous reserve, FUNAI anthropologists investigated the historical-cultural situation of the remaining Aikanã (see the report by Jane Galvão, 1980). It was only in the 1980s that serious initiatives were taken by linguists to study the Aikanã language in detail. In 1984, the linguist Harvey Carlson (deceased in 1994) spent two months with the Aikanã, intending to conduct an extensive study of the language. Although he did not continue this endeavour, Carlson left a collection of tape recordings and field notes at the University of California, Berkeley, which formed the basis of a preliminary study of aspects of the language (Hinton (ed.) 1993). In the 1990s the Brazilian linguist Ione Vasconcelos conducted a study of the language, which resulted in several articles (e.g. 1996, 2005) and

---

9.   Upon first contact in 1977 the Latundê caught influenza and were decimated. Their population decreased by 60% within a year, leaving them more dependent on the Aikanã and the outside world.

a doctoral dissertation (2002). In 2012, Fátima da Silva, a student of the linguist Henri Ramirez, defended a description of the Aikanã lexicon as her MA thesis at the Universidade Federal de Rondônia (Silva 2012).

My own involvement with the Aikanã began in 1994, when I began a study of the neighbouring Kwaza language. Since the Kwaza live among the Aikanã, and because many speakers of Kwaza also speak Aikanã, I had the opportunity to elicit and register some Aikanã data during my research of the Kwaza language. Furthermore, in order to evaluate claims that Kwaza represents a language isolate (Loukotka 1968: 163–164), it was necessary to compare it to the neighbouring languages, among which is Aikanã. One of the results of this comparison was a provisional confirmation of the classification of Aikanã as a language isolate (see van der Voort 2005). In the context of a more general comparative project (see Crevels & van der Voort 2008), I continued to work on Aikanã, and I have built a collection of audio and video recordings and other relevant documents concerning the Aikanã people, in addition to transcriptions and field notes, copies of which were deposited at the archive of the Museu Paraense Emílio Goeldi in Belém.

6.   The modern history and present situation of the Kwaza

Kwaza is spoken by around 25 individuals belonging to three families. One family, of which only the oldest generation speaks the language, lives in the indigenous reserve Terra Indígena Tubarão-Latundê, another family lives mainly in the nearby town of Chupinguaia, and a third family of Kwaza speakers lives in the Terra Indígena Kwazá do Rio São Pedro. The latter two families are related, and the language is still acquired by the youngest generation. The last ethnic Kwaza who was born in traditional times died in April 2008. Unfortunately, he did not pass on some of his knowledge of indigenous culture, and certain information can only be obtained from elderly Aikanã.

The Kwaza originally lived on the Rio São Pedro, the Rio Taboca, and probably other western headwaters of the Pimenta Bueno River. Their traditional neighbours were the Aikanã, Kanoê, Salamãi, Kepkiriwat (extinct TUPIAN), and, at a greater distance, the Latundê and the Mekens and perhaps also the Tupari and the Arikapu who lived across the great savannah area. Relationships with their neighbours were not always friendly.

The Kwaza were usually encountered in the literature as the *Koaiá*, and, very rarely, as *Arara*. Presently, the autodenomination *Kwaza* is used increasingly by Indians and Westerners alike. The Kwaza were mentioned for the first time on the 1913 manuscript map of the Comissão Rondon as *Coaiás*. The Kepkiriwat leader from whom the information was obtained located them on the São Pedro and

Taboca rivers (Rondon 1916; Rondon & Faria 1948). According to the Nambik-wara, they were cannibals and they, as well as other groups of the region, feared them (Anonymous 1916: 342).

The first albeit fragmentary account of the Kwaza people, with photographs and a sample of their language, resulted from the encounter of Lévi-Strauss (1955: 382, 1994) and Castro Faria (2001: 148ff.) with the "Mondé" (who turned out to be Salamãi) in 1938. Apparently, several Kwaza individuals either lived among the Salamãi or were there on a visit. Neither Lévi-Strauss nor Castro Faria identified them as Kwaza, but Lévi-Strauss' word list (1995 [1938]), taken from a boy originating from the São Pedro River, leaves no doubt about the language (Loukotka 1963: 12–13; van der Voort 2004: 30–33). Furthermore, a Kwaza woman by the name Makytxa was recognised in some of the photographs in Castro Faria (2001: 151).

At the time of Lévi-Strauss' visit, the Kwaza were probably already decimated, both by exogenous diseases and, according to the Aikanã, by warfare. Even though their language shows many likely traits of contact with other languages and their culture clearly belongs to the Marico cultural complex, they apparently did not maintain very friendly relationships with other groups during the 20th century. In the 1960s they formed only one or two families living among the Aikanã at the mouth of the Tanaru River, and in 1973, they moved together with the Aikanã to the reserve Terra Indígena Tubarão-Latundê. In 2000, a small reserve was demarcated for the Kwaza on their original lands, the Terra Indígena Kwazá do Rio São Pedro. The sparse early documents in combination with modern documentation and analysis of their language contributed to the judicial evidence supporting the land claim (see van der Voort 2007: 253–254, 2008c).

7.   Previous research and documentation of Kwaza language and culture

The Kwaza are probably the least recognised group in the anthropological and linguistic literature about the region before 1995. After Rondon, no direct reference was made to them until the 1940s, when Estanislau Zack included a 220-word list in his report on the S.P.I. expedition (1943), and Victor Dequech mentioned them in his report and on a survey map (1942, 1943). Thereafter, there is again a long period of silence, and the Kwaza were generally presumed extinct until 1984, when the linguist Harvey Carlson recorded a short word list of Kwaza during his fieldwork with the Aikanã. In 1994, I started to study and document the Kwaza language in a five-year project that resulted in an extensive language description (van der Voort 2004) and an areal-comparative study (van der Voort 2005), among other things. Technical and other limitations at the time prevented

me from building a sophisticated corpus of documentation, but I hope to complement the existing corpus – a copy of which is available at the Museu Paraense Emílio Goeldi – within the current DoBeS project with modern digital audio and video documentation.

8.   Genetic and typological profile of the Aikanã and Kwaza languages

The Aikanã and Kwaza languages are genetic isolates, which means that they are apparently not related to any of the world's languages or linguistic families. As mentioned above, areal diffusion may explain part of the similarities between the languages of the region. There are notable similarities between the Aikanã and Kwaza languages that are also shared with Kanoê (the other isolate of southeastern Rondônia, described by Bacelar 2004). Although there is no compelling evidence that they are genetically related to one another (van der Voort 2005), the possibility of a long-distance genetic relationship cannot be excluded.[10] At any rate, their lexicons are very different, and those etymological correspondences that can be found are usually so close that they must be the result of borrowing. Their phoneme inventories are different and display some rare traits. Aikanã has a voiced interdental fricative consonant /ð/ (written as <z>) that can be nasalised and a close front rounded vowel /y/ (written as <ü>). Kwaza has labial and apico-alveolar implosive consonants /ɓ/ and /ɗ/ (written as <b> and <d>) and relatively many vowel contrasts. Aikanã and Kwaza phonemic nasal contrast in vowels is shared with many languages of the region.

Both languages are morphologically rather complex, especially with regard to the verbs. They are predominantly suffixing languages and are characterised by a certain degree of polysynthesis. Both languages have many classifying and valency changing suffixes. Both languages have a dependent-marking pattern in possessive expressions and both dependent and head-marking patterns for argument relations. Many of these morphosyntactic traits are shared with Kanoê and some with other languages of the region. Some traits are shared only between Aikanã and Kwaza, such as the anticipatory switch reference marking system that indicates whether the subject of the next clause will be different or not from that of the current clause. Both have an alternative head-marking possessive construction

---

10. Although no clear systematic sound correspondences were found in basic vocabulary, recent methodological developments in long-distance genetic comparison may come to change our ideas about their isolate status. Preliminary statistical phylogenetic research based on short standardised word lists suggests that Aikanã and Kwaza have a higher probability of being genetically related to one another than to any other language (Søren Wichmann p.c.).

that concerns only the third person and involves a somewhat similar morpheme (Aikanã *-deri*, Kwaza *-tjate*; see van der Voort 2009b: 355–358, 364–365). The Aikanã future tense construction can only be understood through the Kwaza quotative construction (see van der Voort 2013). In Aikanã, future tense is expressed by the morpheme *-re-*, which is obligatorily preceded by a first person marker whereas the actual subject is expressed by a following person marker, as in example (1):

(1) *hari-txa-ʼre-mia-ē*
     bathe-1PL-FUT-2PL-DECL
     'You people will take a bath.'

In Kwaza several modalities are indicated by a productive quotative construction in which there are two layers of person and mood inflexions, the first layer expressing the quoted event, and the second layer expressing the event of quoting, as in example (2):

(2) *daʼnĩ hɨhɨrwa-a-ʼni=xa-re*
     still   walk-1PL-EXH=2-INT
     'Are we still going for a walk?' (lit.: 'Do you still say: "Let's walk!"?')

Aikanã does not have a quotative construction similar to Kwaza, but its future construction seems to reflect such a quotative construction. Only if the Aikanã example is interpreted literally as 'You people say: "We will take a bath."' its morphological structure involving a first person marker makes sense. This construction may represent a wider areal feature that includes Chapacuran and perhaps Andean languages (van der Voort 2009a: 279, 2013: 372; see also Everett 2008).

The person marking systems of Aikanã and Kwaza are very different, however. Whereas Kwaza displays an inclusive/exclusive distinction, Aikanã does not. Whereas Kwaza has one single paradigm of person inflexions for all verbs, Aikanã has perhaps ten different verbal inflexion classes with different sets of person markers. Whereas Kwaza person marking is exclusively suffixing, several Aikanã verb classes require prefixation of person markers. In that respect Aikanã shows similarities with Kanoê, which also has multiple verb classes and both prefixing and suffixing person markers. For linguists, language isolates are particularly interesting because they may have properties that are very rare or even unique among the world's languages. So far, Kwaza is the only language in which reduplication determined by morphological boundaries as opposed to phonotactic boundaries – in this case reduplication of bound person markers with different syllable structures – was attested unambiguously (van der Voort 2009a: 270–271). Whereas Kwaza has been described relatively thoroughly, an exhaustive description and analysis of Aikanã is still underway.

9.   The importance and urgency of Aikanã and Kwaza documentation

The southwestern Amazon region harbours an extraordinary concentration of linguistic diversity that includes over 10 language isolates. Three of those, Aikanã, Kanoê and Kwaza, are spoken by neighbouring ethnic groups in southeastern Rondônia. The isolate status of their languages could mean that they are among the oldest ethnic groups of the region. As made clear in the above sections, the Aikanã and Kwaza have been relatively neglected in the standard literature on the indigenous peoples of southeastern Rondônia. Furthermore, their cultural and linguistic documentation according to modern standards is rather incomplete and highly urgent. Finally, although the Kwaza language (25 speakers) has been relatively thoroughly described and analysed,[11] the work on Aikanã (175 speakers) is still ongoing.

Due to the very small size of the Kwaza community (which actually cannot be regarded as a distinct community anymore at all) and the passing away of one of the last knowledgeable bearers of traditional culture in 2008, the possibilities of documentation of Kwaza are necessarily limited. However, the process of demarcation of the São Pedro reserve has raised the people's interest in revitalisation of language and culture. Furthermore, indirect reports on traditional Kwaza culture are still available from descendants of deceased Kwaza elders, from certain knowledgeable Aikanã elders and from sparse references in published and unpublished sources. Remaining knowledge of Kwaza oral traditions can still be videotaped, in both the Tubarão-Latundê and São Pedro reserves. What remains of the material culture can still be documented in the São Pedro reserve. So far, the results of these ongoing efforts, which include extensive audio and video recordings, are in the process of being analysed and will be presented in a structured digital archive that will be accessible online for the indigenous communities. These efforts are bound to benefit the indigenous communities, who are increasingly interested in language and culture preservation.

The main focus of current studies concerns the documentation of Aikanã language and culture. The community is of considerable size and all generations are still represented. The language is still acquired by many young people, and there are even second language speakers of Aikanã. However, the possibilities for endogamous marriages have been practically exhausted and most new marriages are between Aikanã and other ethnicities, especially Latundê, Rikbaktsa (MACRO-JÊ),

---

11.   Which does not guarantee that the language is properly documented according to modern standards, with ample high-quality and systematically catalogued and stored audio and video recordings of a representatively wide range of linguistic and cultural events. See Section 10 for an inventory of different types of documentation that the present project aims at.

Terena (ARAWAKAN), Kwaza and Brazilian, in which nowadays Portuguese usually is the family language. Furthermore, the Aikanã language suffers from an 'ecological' decline (i.e. with regard to domains of usage) and a consequent lack of prestige. Since the elderly bearers of language and culture have been passing away one after another, especially in recent years, knowledge of the wonderfully rich oral traditions and the specialised registers of language use that managed to survive in the traditional tales are under serious threat. The language is still spoken fluently on a daily basis by most Aikanã, but few people can still recount traditional tales in their extended form, and few can interpret magic discourse, respectful style of speech and other specialised registers embedded in the stories. For example, respectful speech in Aikanã and its grammatical properties are only used among a handful of elders and are not well understood anymore by those who were born after the 1950s. By way of contrast, the increasing interaction with Western culture has led to an increase of contexts of language use from which Aikanã is excluded.

In addition to the endangered varieties of speech dominated by the elderly, it is also important to pay attention to children's and child-directed speech while this is still possible. The disappearance of the speech of the elderly can be predicted with certainty. In contrast, the vanishing of native first language acquisition is often discovered only after the fact. As mentioned above, representatives of the youngest generations still acquire Aikanã and (surprisingly even) Kwaza as first languages. Some documentation of children's speech exists for the Kwaza language, and attempts are being made to document children's and child-directed speech for Aikanã.

Aikanã language and culture are especially intertwined with traditional music. Although their musical instruments such as flutes, horns and rattles unmistakenly bear the characteristics of the general musical culture of the region, as described especially by Snethlage (1939), they play a central role in myths about the origin of the Aikanã people. The melodies played on the sacred *Purikai* flutes, which were used in now extinct initiation rituals, represented words, but not the lyrics of any established songs. In the 1990s there was still a handful of elderly men who could play the flutes[12] and could interpret these words. Today there is only one man left who plays the *Purikai* flute. Furthermore, collective singing has almost disappeared. At festivities, people sometimes still sing traditional Aikanã songs, which requires a lead singer and the group dancing together in a circle. The lyrics are in Aikanã, but their meaning and sometimes even the specific words are not readily interpretable. There are two female lead singers left who can actually

---

12. In 1997 I recorded the last three players together on VHS. The recording was recently edited at the Museu Goeldi by van der Voort with Markley (2010) and distributed on DVD among the Aikanã.

understand and explain the meaning of the songs. Traditional music became obsolete in the 1950s with the decline of the culture in which it was embedded,[13] but it managed to survive as a sentimental reminder of the past for the elderly. Since the 1980s, when the possession of radios and cassette players became common and when Aikanã music was actively rejected by the protestant missionaries, it gradually lost popularity. Nowadays many young people are mainly disinterested in traditional music, preferring to listen to Brazilian *sertanejo* and *forró*; moreover, various descendants of the traditional musicians find it painful to hear and see the Aikanã music performed since it reminds them of their deceased parents. It seems that, at least for the latter, the traditional music is not associated anymore with a bygone culture but rather with deceased individuals. Nevertheless, there are still young Aikanã who appreciate the traditional music, and although their interest may not yet be strong enough for a revivalist movement, there are signs of hope. The indigenous schoolteachers have lately been using recordings to analyse and teach traditional singing to the children, and some young men have requested and directed documentation of flute music with the specific goal of learning to play themselves.

Although traditional Aikanã music and oral culture form one of the central issues in the documentation project, various surviving aspects of the original material and intellectual cultures, or the memory of those or active reconstructions, are being documented as well. All types of family relationships are still represented,[14] hunting practices such as charming and calling game, bringing home the kill in makeshift baskets out of leaves tied up with *embira*,[15] etc. can still be documented (such traditions are particularly threatened because game is low due to increasing deforestation). People still plant traditional crops on swidden tracts in the forest, although the communal approach is practically gone. Bows and arrows have been long out of use, but some people can still make them, just like flutes, hats, and ornaments. Adapted versions of traditional body ornaments ('artesanato') are made both for personal use and for sale in town. A traditional house can still be made. Although shamanic practices are now virtually gone since the last Aikanã shaman Capitão Arui Uhunei died in 1985, it was possible to interview the shaman's widow and document several healing techniques, as well as

---

13. One informant, born in 1954, remembered the last time that Aikanã music was played in its original function and context, which he said was at the end of the 1950s.

14. To the contrary, among the Kwaza, where due to the small size and fragmentation of the community, specific family roles represented by certain kinship terms are not fulfilled anymore by actual persons.

15. Inner bark of certain trees used as rope.

shamanic songs, that no one knew were still remembered. The widow's children had never heard them before. Finally, the traditional game of head-ball is only remembered by the elderly and has been replaced by Brazilian football.

The elderly remember many places of traditional importance outside of the reserve, even though the landscape has been unrecognisably changed by the farmers (see Comunidade Aikanã et al. 2010). In 2009 a small expedition was organised with some Aikanã to map places along the Pimenta Bueno River and to locate an untouched cave with rock carvings for which the Aikanã have a mythological explanation involving the traditional flutes. Archeological research is called for here, both for prehistory (e.g. the cave) and post-contact history (e.g. the last settlement at Tanaru, the Tubarão-Latundê reserve). Such research is actually very urgent, since e.g. the nearby Cascata falls, which play a special role both in mythological prehistory and in contact history, were recently destroyed in a hydroelectric dam project, and more projects along the Pimenta Bueno River are planned. Officially, archeological patrimony is protected by law, and there is increasing interest in Rondônia by the Brazilian scholarly archeological community. The DoBeS project's growing corpus of documentation includes the native Aikanã and Kwaza perspectives on archeological remains, which calls for the investigation and preservation of that patrimony.

There is an urgent need for serious cultural anthropological study of the Aikanã, not only investigating the culture of the past, but also the changing culture of the present. Since both the environmental and cultural transitions in the southwestern Amazon are happening very fast and have gone very far already, documentation of Aikanã culture is extremely urgent. Furthermore, there is a need for an integrated ethnohistory of southeastern Rondônia, combining existing literature with interviews with the elderly persons. Aspects of the region's ethnohistory are discussed in work on Kwaza, and some reports on Aikanã traditions were documented from now-deceased individuals. Furthermore, elderly persons are helping to interpret information from the early documents by Dequech and others.[16] Consequently, the digital archive resulting from the documentation project incorporates these relevant older documents.

Obviously, other, neighbouring languages and cultures in southeastern Rondônia also require documentation, but the present study is already ambitious enough as it is for its practical limits. Furthermore, the Mekens and Salamãi heritages are being documented and studied by Vilacy Galucio and Denny Moore, respectively, at the Museu Goeldi in Belém; the Kanoê isolate language and culture

---

16. The late Capitão Pedro Aikanã was identified on Dequech's photographs from the early 1940s. Lisa Grund's 2013 photograph of Capitão Pedro, who passed away in January 2014, won the first place in a Max Planck photo contest.

are being documented and studied by Laércio Bacelar in a Museu do Índio project; the study of Latundê has been an integral part of a descriptive, comparative and ethnohistorical project led by Leo Wetzels of the Vrije Universiteit in Amsterdam; the Akuntsu language is presently under study by Carolina Aragon at the University of Utah (under the supervision of Lyle Campbell, University of Hawai'i); and the Jabuti languages have been the subject of studies by myself and others. But the most blatant gap in documentation in southeastern Rondônia, one of the most interesting and besieged regions of Amazonia, still concerns Aikanã and Kwaza. So in our ongoing project, we focus on a complete integrated multidisciplinary documentation of the Aikanã and Kwaza languages and their social, ethnohistorical and ecological contexts. The end results should be accessible as an online historical record. It should be the basis for educational material, for language and culture revitalisation, and for scientific research in various disciplines.

The Aikanã and Kwaza are increasingly aware of the cultural loss involved in the changes in their world and of the value of their (in the case of Aikanã, still very healthy) native languages (see also van der Voort 2007: 256–258). They had always been happy to receive copies of older documents, photos and traditional music recordings. When cassette tape recorders became common, they sometimes made recordings themselves, but most of these got lost or deteriorated eventually. They were quite impressed by the modern possibilities of documentation as exemplified by the online DoBeS Aweti archive, when I was permitted to demonstrate parts of it in 2009 on a computer in Vilhena, and they requested my help in preserving their own knowledge. Since then, a considerable corpus has been built, and in 2013 Cândida Aikanã was hired as an indigenous research and documentation assistant. Her fascination with her language and culture, her dedication and talent, and her standing in the community made her an ideal candidate. She was introduced to the equipment and recording procedures and has been playing a crucial role in all aspects of the project; interviewing, documentation, transcription, translation, and editing.

10. Concluding remarks and preliminary results

This article reports primarily on an ongoing documentation and preservation project. As demonstrated in the previous sections, there are compelling reasons to focus on the Aikanã and Kwaza languages and their ethnohistorical context. Both languages are isolates, one of which has not yet been satisfactorily described, and both are poorly documented. The cultures and histories of their speakers have hardly been investigated and integrated into the overall ethnolinguistic picture of southeastern Rondônia. The urgency of filling this gap is high because

the languages are seriously endangered, due ultimately to the ongoing ecological destruction of Rondônia and the encroachment of Western culture.

The situation of the languages of the region resembles in many ways the one described recently by Gruzdeva (cf. Abstract LAUD 2014) for the island of Sakhalin, northeastern Siberia. Similar to the role of the Russian language on Sakhalin, the Portuguese language is omnipresent among the indigenous peoples in Rondônia. Fortunately, there is increasing awareness among the indigenous peoples, as well as among outsiders, of the importance of the endangered linguistic and cultural heritage of the region. However, favourable initiatives to strengthen the native languages are often not backed by the availability of competent teachers and educational material. Furthermore, native speakers and culturally knowledgeable elders are sometimes impatient with the younger generation and, rather than teaching what they know, criticise what the younger generation does not know. Also, there is a lack of knowledge and expertise at the local universities with regard to the impressive indigenous linguistic wealth in the region.

One of the things that seems to be different from the situation on Sakhalin is the phenomenon that most indigenous languages in southeastern Rondônia continue to be transmitted to the younger generations, sometimes at family level only, even when there is, as in the case of Kwaza, no linguistic community to speak of any longer. If the languages and cultures of southeastern Rondônia are granted support and respect, both from outsiders and the indigenous peoples themselves, and are allowed to adapt to the ever-changing circumstances, they should be able to survive into the next century. As linguists interested in indigenous language and culture, it is scientifically responsible to dedicate a considerable part of our resources to support the survival of that heritage. An accessible permanent record of language and culture represents an important tool for our scientific research as well as a repository of cultural knowledge for the indigenous peoples themselves.

The documentation work within the present project involves the audio and video recording of linguistic, cultural and historical data of the Aikanã and Kwaza peoples. The indigenous communities have co-determined the kinds of data to be documented and are playing a decisive role in questions of accessibility. The corpus so far consists of the following (interconnected) kinds of data:

–   texts of different kinds
    –   traditional mythology
    –   direct and indirect personal histories (from both genders of all ages)
    –   conversational dialogue (from both genders of all ages)
    –   procedural texts

- elicited linguistic data
  - phonological minimal pairs
  - word lists (basic and specific, e.g. ethnobiological)
  - specific linguistic phenomena (e.g. demonstratives, classifiers, etc.)
- traditional music and dancing
  - shamanic singing
  - festive singing
  - sacred flutes
  - festive flutes
- aspects of culture
  - daily practices (hunting, cooking, building, etc.)
  - environmental knowledge (flora, fauna, climate, geography)
  - visits to culturally or historically relevant places in traditional territory
  - interviews (about culture, history, music, religion, ecology, etc.)
- lexical and grammatical databases
  - comprehensive Aikanã database under construction
  - existing Kwaza database

Most of the data are being recorded both in audio and video format and are being transcribed, interlinearly glossed, and translated into Portuguese and English. They are annotated according to standard data and metadata management practices involving the appropriate hardware and software, such as ELAN and IMDI.

The resulting Aikanã corpus, the Kwaza corpus and existing database, and minor Salamãi and Latundê corpora are being digitally stored in the digital(ised) language documentation archives of the Museu Paraense Emílio Goeldi in Belém and the Max Planck Institute for Psycholinguistics in Nijmegen. These archives will be available online, in differentiated levels of accessibility, for the speakers' communities and for third parties. These corpora will serve the indigenous communities as a repository of endangered traditional knowledge and as a basis for preservation and revitalisation of language and traditional aspects of culture. From it, data can be extracted and made available in the form of audio and video discs (which is increasingly common in the indigenous communities) and printed work for the development of educational material, for example. Various disks of important non-confidential recordings have already been edited and distributed in the community by the team. They are enormously popular and are also used in the indigenous schools. The corpus will also serve the scientific community as a unique archive and research database, especially with regard to the ethnohistory and ethnobiology of the region, to the description of Aikanã lexicon and grammar, and to theoretical and long-distance comparative linguistic issues.

Previously existing data, when relevant, are being integrated into the corpus. This concerns especially mythological texts, music and certain linguistic documentation. Such older material is at times the result of provisional exploratory documentation on audio and requires re-recording in a more controlled and complete manner. Both the recording and the analysis of all material is realised in close cooperation with indigenous assistants, who are trained thoroughly in the hopes that they will be able to carry out their own documentary projects in the future.

Acknowledgements

Abbreviations

1	first person	FUT	future
2	second person	INT	interrogative
DECL	declarative	PL	plural
EXH	exhortative		

References

Albert, C.P. 1964. Americo Casara: Conquistador pacifique de l'Amazonie. *Lettre d'Amazonie* 7: 7–11, 8: 6–13.

Alves, P.M. 2004. O léxico do Tuparí: Proposta de um dicionário bilíngüe. PhD dissertation, Universidade Estadual Paulista. Júlio de Mesquita.

Anonymous. 1916. *Missão Rondon: Apontamentos sobre os trabalhos realizados pela Commissão de Linhas Telegraphicas Estrategicas de Matto-Grosso ao Amazonas, sob a direcção do coronel de engenharia Cândido Mariano da Silva Rondon, de 1907 a 1915*. Rio de Janeiro: Typ. do Jornal do Commercio, de Rodrigues & C.

Aragon, C.C. 2008. Fonologia e aspectos morfológicos e sintáticos da língua Akuntsú. MA thesis, Universidade de Brasília.

Bacelar, L.N. 2003. Kanoê. In *Enciclopédia dos povos indígenas no Brasil*, B. & F. Ricardo (eds). São Paulo: Instituto Socioambiental. <http://pib.socioambiental.org/pt/povo/kanoe>

Bacelar, L.N. 2004. Gramática da língua Kanoê. PhD dissertation, Katholieke Universiteit Nijmegen. <http://webdoc.ubn.kun.nl/mono/b/bacelar_l/gramdalik.pdf>

Becker-Donner, E. 1955. Notizen über einige Stämme an den rechten Zuflüssen des Rio Guaporé. *Archiv für Völkerkunde* X: 275–343. Wien: Wilhelm Braumüller Universitäts-Verlag.

Bontkes, W. 1968. *Kasupá. Formulário dos vocabulários padrões para estudos comparativos preliminares nas línguas indígenas Brasileiras*. Rio de Janeiro: Museu Nacional 19.

Braga, A. de O. 2005. Aspects morphosyntaxiques de la langue makurap. PhD dissertation, Université Toulouse II.

Brijnen, H. & Adelaar, W.F.H. 2010. *Amazonian languages*. Leiden: Universiteitsbibliotheek. <https://digihum.leidenuniv.nl/amazonianlanguages>

Caspar, F. 1953. Ein Kulturareal im Hinterland der Flüsse Guaporé und Machado (Westbrasilien), dargestellt nach unveröffentlichten und anderen wenig bekannten Quellen, mit besonderer Berücksichtigung der Nahrungs- und Genussmittel. PhD dissertation, Universität Hamburg.

Caspar, F. 1975. *Die Tuparí: Ein Indianerstamm in Westbrasilien* [Monographien zur Völkerkunde VII]. Berlin: Walter de Gruyter.

de Castro, T.P.V. 2012. Djeoromitxí: Notes on phonology and simple noun phrase structure. MA thesis, University of Texas at Austin.

de Castro Faria, L. 2001. *Um outro olhar: Diário da expedição à Serra do Norte*. Rio de Janeiro: Ouro sobre Azul.

Comunidade Aikanã, Doelman, W. & van der Voort, H. 2010. *Mapa etno-histórico dos Aikanã*. Amsterdam: Own.

Crevels, M. 2012. Language endangerment in South America: The clock is ticking. In *The indigenous languages of South America: A comprehensive guide*, L. Campbell & V. Grondona (eds), 167–233. Berlin: De Gruyter Mouton.

Crevels, M. & van der Voort, H. 2008. The Guaporé-Mamoré region as a linguistic area. In *From linguistic areas to areal linguistics* [Studies in Language Companion Series 90], P. Muysken (ed.), 151–179. Amsterdam: John Benjamins. doi:10.1075/slcs.90.04cre

Dequech, V. 1942. Trabalhos da commissão do Urucumacuan nos rios Apediá e Corumbiára. Ms, map.

Dequech, V. 1943. *Commissão para o estudo das jazidas auriferas do Urucumacuan: Relatorio apresentado pelo eng. de minas e civil Victor Dequech*, 38 ff. Rio de Janeiro: Divisão de Fomento da Produção Mineral.

Duran, I.R. 2000. Descrição fonológica e lexical do dialeto Kaw Tayo da língua Moré. MA thesis, Universidade Federal de Rondônia.

Everett, C. 2006. Gestural, perceptual, and conceptual patterns in Karitiana. PhD dissertation, Rice University.

Everett, D. 2008. Wari' intentional state constructions. In *Investigations of the syntax-semantics-pragmatics interface* [Studies in Language Companion Series 105], R. Van Valin Jr. (ed.), 381–409. Amsterdam/Philadelphia: John Benjamins. doi:10.1075/slcs.105.27eve

Everett, D. & Kern, B. 1997. *Wari': The Pacaas Novos language of Western Brazil*. London: Routledge.

Fawcett, P.H. 1915. Bolivian exploration, 1913–1914. *Geographical Journal* 45: 219–228. doi:10.2307/1779796

da Fonseca, J.S. 1880–1881. *Viagem ao redor do Brasil, 1875–1878*. 2. Rio de Janeiro: Pinheiro.

de França, M.C.V. 2002. Aspectos da fonologia lexical e pós lexical da língua Oro Towaťi (Oro Win). PhD dissertation, Universidade Federal de Rondônia.

Gabas Jr., N. 1999. A grammar of Karo, Tupí (Brazil). PhD dissertation, University of California at Santa Barbara.

Galucio, A.V. 2001. The morphosyntax of Mekens (Tupi). PhD dissertation. University of Chicago.

Galucio, A.V. 2005. Puruborá: Notas etnográficas e lingüísticas recentes. *Boletim do Museu Paraense Emílio Goeldi (Ciências Humanas)* 1(2): 159–192.

Galucio, A.V. 2006. *Narrativas tradicionais Sakurabiat mayãp ebõ*. Belém: Museu Paraense Emílio Goeldi.

Galvão, J.L.F. 1980. *Relatório de identificação de área indígena dos Aikaná do Território Federal de Rondônia*, 142 ff. Rio de Janeiro: FUNAI, Museu do Índio.

Gomes, M. Pereira. 1991 [1988]. *Os Índios e o Brasil*. Petrópolis: Vozes.

Grund, L.K. 2013. [untitled winning photo]. Photo competition A Picture of Science – Taal in Beeld. Max Planck Institute, Radboud Universiteit, & Bibliotheek Mariënburg, Nijmegen. <http://pictureofscience.ruhosting.nl/wp-content/uploads/welcome_flyer.pdf>

Gruzdeva, E. 2014. Indigenous languages of Sakhalin Island: On the verge of extinction. *Abstracts LAUD 2014. 36th International LAUD Symposium*. March 31–April 3, 2014. Landau, Germany.

Guerra, M. de L. 2004. Aspects of Suruí phonology and phonetics. MA thesis, Université Libre de Bruxelles.

Hanke, W. 1956. Beobachtungen über den Stamm der Huari (Rio Corumbiara Brasilien). *Archiv für Völkerkunde* XI: 67–82. Wien: Wilhelm Braumüller Universitäts-Verlag.

Hinton, L. (ed.). 1993. *Aikana modules; A class report on the fieldnotes of Harvey Carlson*, 158ff. Berkeley CA: University of California.

Holtwijk, I. 2006. *Rooksignalen: Op zoek naar de laatste verborgen Indianen in Brazilië*. Amsterdam/Antwerpen: Atlas.

Lévi-Strauss, C. 1948. Tribes of the right bank of the Guaporé River. In *Handbook of South American Indians, 3: The tropical forest tribes*, J.H. Steward (ed.), 370–379. New York NY: Cooper Square Publishers.

Lévi-Strauss, C. 1955. *Tristes Tropiques*. Paris: Librairie Plon.

Lévi-Strauss, C. 1994. *Saudades do Brasil*. São Paulo: Companhia das Letras.

Lévi-Strauss, C. 1995. Untitled Manuscript: Copy of 1938 word list and letter dated 25-12-1995. Paris, 3.

Loukotka, Č. 1963. Documents et vocabulaires inédites de langues et de dialectes sud-américains. *Journal de la Société des Américanistes* LII: 7–60. doi:10.3406/jsa.1963.2001

Loukotka, Č. 1968. *Classification of South American Indian languages*. Los Angeles CA: Latin American Center, University of California.

Maldi Meireles, D. 1989. *Guardiães da fronteira: Rio Guaporé, século XVIII*. Petrópolis: Vozes.

Maldi, D. 1991. O complexo cultural do Marico: Sociedades indígenas dos Rios Branco, Colorado e Mequens, afluentes do Médio Guaporé. *Boletim do Museu Paraense Emílio Goeldi* 7(2): 209–269.

Mindlin, B. 1985. *Nós Paiter: Os suruí de Rondônia*. Petrópolis: Editora Vozes.

Moore, D. 1984. Syntax of the language of the Gavião Indians of Rondônia, Brazil. PhD dissertation, City University of New York.

Monserrat, R.M.F. 2005. Noticia sobre a lingua purubora. In *Novos estudos sobre línguas indígenas*, A. Rodrigues & A.S. Cabral (eds), 19–22. Brasilia: Universidade de Brasilia.

Muysken, P., Hammarström, H., Krasnoukhova, O., Müller, N., Birchall, J., van de Kerke, S., O'Connor, L., Danielsen, S., van Gijn, R. & Saad, G. 2014. *South American Indigenous Language Structures (SAILS) Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://sails.clld.org/>

Nogueira, A.F. de S. 2011. Wayoro ẽmẽto: Fonologia segmental e morfossintaxe verbal. MA thesis, Universidade de São Paulo.

Nordenskiöld, E. 1915. *Forskningar och Äventyr i Sydamerika*. Stockholm: Albert Bonniers Förlag.

Pickering, W. 1968. *Mondé. Formulário dos vocabulários padrões para estudos comparativos preliminares nas línguas indígenas brasileiras*, 5ff. Rio de Janeiro: Museu Nacional.

Pires, N.N. 1992. Estudo da gramática da língua Jeoromitxi (Jabutí) – Aspectos sintáticos das cláusulas matrizes. MA thesis, Universidade Estadual de Campinas.

Price, P.D. 1983. Pareci, Cabixi, Nambiquara. A case study in the western classification of native peoples. *Journal de la Société des Américanistes* 69(1): 129–148. doi:10.3406/jsa.1983.2228

Reel, M. 2010. *The last of the tribe: The epic quest to save a lone man in the Amazon*. New York NY: Scribner.

Reesink, E. 2012. *Allegories of wildness: Three Nambikwara ethnohistories of sociocultural and linguistic change and continuity*. Amsterdam: Rozenberg.

Rivet, P. 1953. La langue Mašubi. *Journal de la Société des Américanistes* XLII: 119–125. doi:10.3406/jsa.1953.2403

Rodrigues, A. & Caspar, F. 1958. *Zur Sprache der Tuparí Indianer (Westbrasilien)*, 95ff.

Rondon, C.M. da S. 1916. *Conferencias realizadas nos dias 5, 7 e 9 de Outubro de 1915 pelo Sr. Coronel Cândido Mariano da Silva Rondon, no Theatro Phenix do Rio de Janeiro sobre trabalhos da Expedição Roosevelt e da Commissão Telegráfica*. Rio de Janeiro: Typ. Leuzinger.

Rondon, C.M. da S. & de Faria, J.B. 1948. *Glossário geral das tribos silvícolas de Mato Grosso e outras da Amazônia e do norte do Brasil*. Rio de Janeiro: Ministério da Agricultura, Conselho Nacional de Proteção aos Índios.

Sampaio, W. B. de A. 2001. As línguas tupi-kawahib: Um estudo sistemático filogenético. PhD dissertation, Universidade Federal de Rondônia.

Silva, M. de F. dos S. de. 2012. Dicionário de raízes da língua Aikanã. MA thesis, Guajará-Mirím: Universidade Federal de Rondônia.

Snethlage, E.H. 1937. *Atiko Y: Meine Erlebnisse bei den Indianern des Guaporé*. Berlin: Klinkhardt & Biermann.

Snethlage, E.H. 1939. Musikinstrumente der Indianer des Guaporégebietes. *Baessler-Archiv, Beiträge zur Völkerkunde*. Berlin: Dietrich Reimer – Andrews & Steiner.

Storto, L. 1999. Aspects of a Karitiana grammar. PhD dissertation, MIT.

Telles, S. 2002. Fonologia e gramática Latundê/Lakondê. PhD dissertation, Vrije Universiteit Amsterdam.

Vasconcelos, I.P. 1996. Algumas considerações sobre a morfologia Aikanã. *ABRALIN, Boletim da Associação Brasileira de Lingüística* 19: 71–78.

Vasconcelos, I.P. 2002. Aspectos fonológicos e morfofonológicos da língua Aikanã. PhD dissertation, Universidade Federal de Alagoas.

Vasconcelos, I.P. 2005. Aikanã. In *Enciclopédia dos povos indígenas*, B. & F. Ricardo (eds). São Paulo: Instituto Socioambiental. <http://pib.socioambiental.org/pt/povo/aikana>

van der Meer, T. 1982. Fonologia da língua Suruí. MA thesis, Universidade Estadual de Campinas.

van der Voort, H. 2004. *A grammar of Kwaza*. Berlin: Mouton de Gruyter.
doi: 10.1515/9783110197280

van der Voort, H. 2005. Kwaza in a comparative perspective. *International Journal of American Linguistics* 71(4): 365–412. doi: 10.1086/501245

van der Voort, H. 2007. Theoretical and social implications of language documentation and description on the eve of destruction in Rondônia. In *Language documentation & linguistic theory: 75 years of linguistics at SOAS, 5 years of Endangered Languages Project*, P.K. Austin, O. Bond & D. Nathan (eds), 251–259. London: School of Oriental and African Studies, University of London. <http://www.hrelp.org/publications/ldlt/papers/ldlt-proceedings.html>

van der Voort, H. 2008a. Arikapú. In *Enciclopédia dos povos indígenas*, B. & F. Ricardo (eds). São Paulo: Instituto Socioambiental. <http://pib.socioambiental.org/pt/povo/arikapu>

van der Voort, H. 2008b. Djeoromitxí. In *Enciclopédia dos povos indígenas*, B. & F. Ricardo (eds). São Paulo: Instituto Socioambiental.<http://pib.socioambiental.org/pt/povo/djeoromitxi>

van der Voort, H. 2008c. Claude Lévi-Strauss' møde med Kwazá-indianere i 1938 og sprogvidenskabens betydning for samfundet. *Jordens Folk* 43(3): 34–43.

van der Voort, H. 2009a. Reduplication and repetition of person markers in Guaporé isolates. *Morphology* 19(2): 263–286. <http://www.springerlink.com/content/1545301353748536/fulltext.pdf> doi: 10.1007/s11525-009-9143-8

van der Voort, H. 2009b. Possessive expressions in the Southwestern Amazon. In *The expression of possession*, W.B. McGregor (ed.), 343–388. Berlin: Mouton de Gruyter.

van der Voort, H., Arikapú, M., Arikapú, N. & Alves, A.C.F. 2010. *Vocabulário Arikapú-Português*. Monografias de Etnolingüística. <http://www.etnolinguistica.org/mono:1>

van der Voort, H. 2012. Whatever happened to Mashubi? Taking a new look at Fawcett's vocabulary. *Cadernos de Etnolingüística* 4(1): 1–20. <http://www.etnolinguistica.org/issue:vol4n1>

van der Voort, H. 2013. Fala fictícia fossilizada: O tempo futuro em Aikanã. *Boletim do Museu Paraense Emílio Goeldi (Ciências Humanas)* 8(2): 359–377. <http://www.scielo.br/pdf/bgoeldi/v8n2/09.pdf> doi: 10.1590/S1981-81222013000200009

van der Voort, H. & Markley, S. 2010. *A música sagrada dos Aikanã: As flautas Purikai*. Belém: Museu Paraense Emílio Goeldi.

Wichmann, S. 2009. Personal communication.

Zack, E. 1943. *Turma de exploração no Oéste de Mato Grosso: Relatorio; Vocabulario das tribus Massacá, Salamãin, Coaiá e Canoê*, 19ff. Rio de Janeiro: Archive of the Museu do Índio.

Metaphors of an endangered forest people, the Yanomae (N. Brazil)

Gale Goodwin Gómez
Rhode Island College

Among the Yanomae-speakers of Brazilian Amazonia the close relationship between language, culture and the environment is revealed in the use of metaphor, metonymy, and euphemism in a wide range of words and phrases. For example, euphemistic metonymy demonstrates how *not* speaking of the dead can provide insights into Yanomami culture as the deceased are referred to by specific items of material culture that identify each individual's role in traditional Yanomami society. This paper provides abundant examples from field data to demonstrate how the traditional culture and rain forest lifestyle are reflected in images created by the Yanomae language. Deeply rooted in rain forest culture, the language will continue to thrive only as long as Yanomami lands remain protected from outside exploitation.

Keywords: Amazonia, Brazil, endangered, euphemism, metaphor, metonymy, rain forest environment, Yanomae, Yanomami

1. Introduction

Metaphor involves looking at a concept from another perspective, extending the meaning of one concept to another. As Lakoff and Johnson define it, "The essence of metaphor is understanding and experiencing one kind of thing in terms of another" (1980: 5). For example, in English a classic metaphor views *time* as a commodity that can be spent, saved, or wasted, like money. Among the Yanomae-speaking people of northern Brazil, the use of metaphor is especially common for neologisms and taboo topics, such as death, as well as in orientational and temporal expressions. This paper examines the use of metaphor and two related phenomena – metonymy and euphemism – in the Yanomae language to demonstrate the close relationship between language, culture and the environment among one

group of indigenous Amazonians who maintain a traditional rain forest lifestyle, despite the ever-encroaching modern world.

Yanomae-speakers belong to one of four major subgroups of the Yanomami language family, found in the Amazon rain forest of northern Brazil and southern Venezuela. The other linguistic subgroups include Yanomamɨ, Sanɨmá, and Ninam (also called Yanam). The author conducted her doctoral dissertation fieldwork among Ninam-speaking Yanomami (Gomez 1990). No definitive genetic relationship has been established between Yanomami languages and any other South American linguistic group. According to the 2011 Brazilian National Health Foundation (FUNASA) census, an estimated 19,000 Yanomami live in Brazil. Of these Brazilian Yanomami, about 47.5% (or an estimated 9,123 people) speak Yanomae. The present paper focuses on the dialect of Yanomae spoken in the community of Watorikɨ located in the state of Amazonas, Brazil. The linguistic data in the present study were collected during personal fieldwork by the author and in collaboration with French cultural anthropologist Bruce Albert, who has worked with the Yanomami since the 1970s. Specific examples are from the manuscript of a thematic lexicon of Yanomae that is in preparation (Gómez & Albert n.d.).

Linda Chinelo Nkamigbo points out that metaphorical proverbs in Koring (a Niger-Congo language of Nigeria) "resulted from the lifestyle and customs of people, who have had strong bonds with the natural world" (Abstract LAUD 2014:113). This same bond is evident among the Yanomae, whose metaphors reflect the flora and fauna that inhabit their natural rain forest environment. The inhabitants of the community of Watorikɨ are hunter-gatherer horticulturalists, whose survival depends on slash-and-burn cultivation of manioc/cassava, bananas, sweet potatoes and other crops as well as gathering wild forest products and hunting for game, especially monkeys, birds, wild pigs, tapirs and small mammals, using bows and arrows. Men are the hunters, and women plant, harvest, and process garden produce. Their culture is built upon these traditional activities, the practice of shamanism, and an animistic belief in spirits of the natural world.

Sustained contact between the inhabitants of Watorikɨ and outsiders began in the 1990s, manufactured goods became commonplace, and a village school was established. These modern innovations brought new words and concepts to the Yanomae language. Since most speakers were (and still are) largely monolingual, many of the terms for these new items are metaphorical extensions of existing vocabulary that directly relates to the Amazon rain forest and traditional cultural practices and beliefs. This new vocabulary will be the focus of section 5 on neologisms.

Sally Rice points out that metaphor and metonymy in Dene Sųłiné are "ubiquitous in colloquial language and do not pertain to a highly composed genre or

register" (2012: 24); this is quite similar to the Yanomae case described here, and all of the examples are from oral language. Most of the paper deals with traditional metaphors, not resulting from contact with outsiders. These metaphors are grouped into sections according to well-established categories of metaphors observed in many languages, beginning with body-part metaphors in Section 2 and moving to orientational and temporal metaphors in Sections 3 and 4, respectively. Section 6 describes the culturally significant taboo against speaking of the dead and how this is maintained through euphemistic metonymy that references the daily activities of the speakers. In Section 7 concluding remarks link the endangerment of metaphors in Yanomae to the threat to the survival of native Amazonians and their traditional lands. The data were collected in Watoriki village during elicitation and interviews conducted jointly by the author and Bruce Albert, who is fluent in Yanomae and provided the Portuguese translations. Discussions with our Yanomae collaborators confirmed the metaphorical nature of the examples, especially the use of euphemism to refer to deceased individuals that reflects a strong cultural taboo. Many of the examples discussed in this paper appeared in a previous publication (Albert & Gómez 1997) that was designed to be used as a bilingual health manual for visiting medical teams and health workers among a mostly monolingual, non-literate Yanomae-speaking population.

2.   Body-part metaphors

A metaphor is "[a]n expression that has a literal interpretation of application in one domain [and] takes on a figurative meaning in another domain" (Rice 2012: 24). Consequently, the mapping of body-parts onto other physical objects is commonplace in languages throughout the world. In English, for example, references to the *neck* of a bottle, the *eye* of a needle, and the *elbow* of a river are classical examples of metaphors. "The parts of the body are the closest and most immediate things in our physical environment, and are thus more deeply imprinted in our cognition, so it is no wonder that body-parts are the sources of terms for all kinds of more abstract concepts in so many languages" (Deutscher 2005: 139).

In Yanomae, objects that have openings, such as a clay pot in example (1), may be described as having a mouth. Likewise, the interior of a house in example (2) is referred to as its chest, which is the center of family and cultural activities much as the human chest holds the most important organs to keep the body alive and active.

(1) *hapaka kahikɨ*
     clay.pot mouth
     'mouth of a clay pot'

(2) *yahi   parɨkɨ-ha*
     hearth chest-LOC
     'inside the house'

Mountains, hills, and rivers are obvious landmarks, and geographical terms are classic sources of body-part metaphors in many languages. In English we speak of *foot*hills and the *foot* of a mountain; whereas, in Yanomae the base of a hill is its 'anus' *xioka* in (5), and the side of the mountain or hill is its 'chest' *parɨkɨ* in example (4). In example (3) in the direction 'toward the top of a mountain/hill,' one is headed (to use another English metaphor!) toward the 'top of the mountain or hill,' which is metaphorically linked to the top of the human body.

(3) *hehu hwesika=hamɨ*
     hill   top.of.head=DIR
     'toward the top of a mountain/hill'

(4) *hehu parɨkɨ-ha*
     hill   chest-LOC
     'on the side of the mountain/hill'

(5) *hehu xioka-ha*
     hill   anus-LOC
     'at the base of the mountain/hill'

With regard to the upper tributaries of a river, the body part 'head' *he* is employed in example (6) in a similar fashion to its English equivalent to designate the '*head-*waters.' The river bank is its 'lip' *kasi* in example (7) and a tributary stream is its 'arm' *poco* in example (8).

(6) *mãu=u   he-kɨ-ha*
     water=CL.liquid head-PL-LOC
     'at the headwaters of the river'

(7) *mãu=u   kasi-ha*
     water=CL.liquid lip-LOC
     'on the river bank'

(8) *poko=u*
     arm=CL.liquid
     'tributary stream of a river'

As mentioned in the introduction, the Yanomae men hunt forest animals with bows and arrows, hence, the importance of these weapons in their culture. Bows and arrows are handcrafted from natural materials, and the men refer to the parts of the bow, *raha sihi*, and arrow, *xaraka*, using body-part metaphors. The parts of a bow include a 'chest' in example (9), a 'back' in (10), and a 'penis' in (11). An interesting detail of Yanomami men's traditional attire might add to understanding this last metaphor; a man's penis is traditionally tied up with a cotton string that connects to a string around the waist. A loose (adult) penis is considered obscene, culturally taboo. The *raha sihi moxi* is where the cord (like a man's penis string) is attached, but it may also be the pointed shape of the end that is a contributing factor for the metaphor. In contrast, the parts of an arrow include the 'head' *he* at one end in example (12) and the 'umbilicus' *makasi* at the other in (13). It should be noted that the 'head' end of the arrow, *xaraka he*, is the one that has the feathers, which might be seen as a link to the feather headdresses worn by men. Perhaps, the *xaraka makasi* end of the arrow, with its repeated wrapping of cotton thread, calls to mind the tied knot of a newborn's umbilical cord. In any case, it is interesting to speculate on the cultural details that seem to relate to certain body-part metaphors.

(9) *raha=sihi   parɨkɨ*
     bow=CL.bow chest
     'inner side of a bow'

(10) *raha=sihi   yaɨpë*[1]
     bow=CL.bow back
     'outer side of a bow'

(11) *raha=sihi   moxi*
     bow=CL.bow penis
     'end of a bow [where cord is attached]'

(12) *xaraka he*
     arrow   head
     '[feathered] end of an arrow'

(13) *xaraka makasi*
     arrow   umbilicus
     'part of an arrow (where the thread is wound just before the arrow point)'

---

1.   Standard IPA symbols are used in the examples, except for /ë/ which represents the 'schwa' sound /ə/ in Brazilian Yanomami orthography and /x/, which is common usage in Brazilian linguistics, for the standard IPA /ʃ/.

3.   Orientational metaphors

There are two opposing spatial orientations: *koro* and *ora*, each of which has a broad range of meanings used in a variety of semantic domains, including the human body, physical objects (such as a canoe), and a story narrative, but they are not common in temporal expressions. Basically, *koro* refers to 'lower or back' while *ora* is 'upper or front.' By extension, *koro* is used to refer to the *lower part* of the body, the *back* of a canoe, or the *end* of a story; while *ora*, in contrast, is the *upper part* of the body, the *front* of a canoe, or the *beginning* of a story. When combined with the word for 'sky' *hutumosi*, *koro* and *ora* refer to cardinal directions 'east' in (14a) and 'west' in (14b). When accompanied by locative and directional morphemes, their meanings are 'downstream' in (15a) and (16a) and 'upstream' in (15b) and (16b).

(14) a. *hutumosi koro*
          sky   lower/back
          'east'

      b. *hutumosi ora*
          sky   upper/front
          'west'

(15) a. *koro-ha*
          lower/back-LOC
          'downstream (location)'

      b. *ora-ha*
          upper/front-LOC
          'upstream (location)'

(16) a. *koro=hamɨ*
          lower/back=DIR
          'downstream (direction)'

      b. *ora=hamɨ*
          upper/front=DIR
          'upstream (direction)'

These orientational terms may also specify nouns, such as *yano* 'house' in (17) and *aka* 'tongue' in (18) to indicate relative positions: *koro* in (17a) refers to the 'lowest point of a communal house,' while in (18a) it refers to the 'back of the tongue.' In contrast, *ora* refers to the 'highest point of a communal house' in (17b) and the 'front of the tongue' in (18b). When co-occurring with the body part 'teeth' *nakɨ*, the terms *koro* and *ora* specify 'lower' and 'higher' parts of a tooth – the root or the crown, respectively, in (19a) and (19b).

(17) a. *yano   koro*
        house   lower/back
        'lowest point of a communal house (at the base of the exterior wall)'

    b. *yano   ora*
        house   upper/front
        'highest point of a communal house'

(18) a. *aka   koro*
        tongue lower/back
        'back of the tongue'

    b. *aka   ora*
        tongue upper/front
        'front of the tongue'

(19) a. *na-kɨ   koro*
        tooth-PL lower/back
        'root of the teeth'

    b. *na-kɨ   ora*
        tooth-PL upper/front
        'crown of the teeth'

Synonyms *oraka* and *orahi,* which refer to the 'neck,' are undoubtedly related to the orientational morpheme *ora*, since the neck is located in the upper body. The term *oraka* is also used to refer to the long 'neck' of a gourd, the entrance to a bee hive, and the top of a closed communal house. Similarly, the phrases *hapaka ora* 'upper part ('neck') of a clay pot' and *kapixa ora* 'clothing worn on the upper part of the body' continue to employ and extend the meaning of *ora*. As a side note, the term *kapixa* is borrowed from the Portuguese *camisa* 'shirt'; traditional Yanomami attire for both men and women is the absence of clothing, except for the penis strings for men and fringed cotton genital coverings for women. Body paint designs are favored over clothing in the hot tropical climate.

4. Temporal metaphors

Deutscher claims that in many languages "no two domains are more intimately linked than space and time…we invariably speak of time in terms of space, and this reflects the fact that we *think* of time in terms of space" (2005: 134), and English demonstrates this with examples like "at the *end* of the year" and "the last *half* of summer." This does not seem to be the case in Yanomae, where the temporal clitic *tëhë* is postposed to noun phrases that describe particular events or characteristics associated with the dry or rainy seasons. For example, favorite rain forest

foods – *paxo wīte* 'fat monkey (meat)' in (20a) or *raxa* 'peach palm (fruits)' in (20b) – are abundant during particular times of the year, and the water level of the rivers is correspondingly higher or lower at specific times, as described in examples (21a) and (21b). The time-related metaphors in these examples highlight the close link between the language and observable events associated with the tropical rain forest.

(20) a. *paxo   wīte=tëhë*
          monkey fat=TEMP
          'fat monkey time (= at the end of the rainy season)

     b. *raxa=tëhë*
          peach.palm=TEMP
          'peach palm (harvest) time' (= summer / dry season)

(21) a. *maũ=u   ōki-o=tëhë*
          water=CL.liquid deep-STATIVE=TEMP
          'high water time' [lit. 'when the river is deep']

     b. *maũ=u   wehe-o=tëhë*
          water=CL.liquid dry-STATIVE=TEMP
          'low water time' [lit. 'when the river is dry']

To further specify a point in time during the evening or night hours (*titi tëhë* 'dark time'), quantifiers that relate to human life stages – *oxe* 'young' in (22a), *hiya* 'young boy' in (22b), and *pata* 'old' in (22c) – are applied to the night. To indicate the midpoint, or midnight, two synonymous expressions are possible. The first (23a) uses the body part, *hriki* 'back,' designating the middle or center of the body, as a metaphor for the middle of the night. The second example (23b) employs a spatial term, *miamo* 'middle,' that refers to the middle or center of a space, which when applied to the temporal duration 'night' produces the meaning 'midnight.' This is one example that links spatial and temporal domains.

(22) a. *titi   oxe-o=tëhë*
          dark   young-STATIVE =TEMP
          'at the beginning of the night' (compare Eng. 'the night is still young')

     b. *titi   hiya-o=tëhë*
          dark   young.boy-STATIVE=TEMP
          'early in the night'

     c. *titi   pata-o=tëhë*
          dark   old-STATIVE=TEMP
          'late at night' [lit. 'when the dark (time) is old']

(23)   a.   *titɨ*   *hrɨk=tëhë*
           dark   back=TEMP
           'at midnight'
     b.   *titɨ*   *mɨamo=tëhë*
           dark   middle=TEMP
           'at midnight'

5.   Neologisms

The physical environment and the culture within which a language thrives and changes provide old meanings that can be extended to new concepts. Among the neologisms found in Yanomae, new vocabulary for modern, manufactured goods provides a clear picture of the speakers' rain forest framework within which these new terms are situated. Examples (24), (25) and (26) show how the referents of objects from the natural environment: a thorn, a gourd, and a woven sac-like recipient (compare *koxikɨ pesi* 'sac-like spider web' and *kaxapë pesi* 'caterpillar's cocoon') are extended to a needle or pin or water bottle.

(24)   *misi=a*
       thorn=SG
       'needle or pin'

(25)   *napë*   *horokoto=e*
       outsider   gourd=POSS
       'bottle' [lit. 'outsider's gourd (used for water)]

(26)   *maũ=u*   *pesi*
       water=CL.liquid   woven.sac
       'bottle of water' (especially bottled mineral water)

Like the examples presented in Section 2, names for new manufactured items frequently employ the same strategy of applying human or animal body-parts to designate new objects. The shape of a spoon resembles a tongue introduced by (or used by) an outsider in example (27) and a machete in example (28) brings to mind a large metal tongue. Example (29) involves the transfer of the name of a traditional knife *thomɨ nakɨ*, which is actually made of agouti teeth, to the new metal object that serves the same purpose.

(27)   *napë*   *aka*
       outsider   tongue
       'spoon'

(28) *poo   aka*
      metal tongue
      'machete'

(29) *thomɨ na-kɨ*
      agouti tooth-PL
      'knife'

Examples (30) through (36) link the body parts of specific rain forest animals (including a mythological being in (35)) to manufactured objects, especially tools, that physically resemble parts of these animals. As hunters are careful observers of their rain forest homeland, the Yanomae speakers immediately associate new objects with others already familiar to them. Because metal was previously unknown to Amazonian peoples and is stronger and more resistant than other known materials, metal tools and objects are highly prized and sought after by the Yanomami. Terms for such useful innovations from outsiders were quickly incorporated into the culture.

(30) *yawere nahasi-kɨ*
      sloth   claw-PL
      'fork' [lit. 'sloth claws']

(31) *hewe yõpa=si-kɨ*
      bat   wing=CL.thin-PL
      'umbrella' [lit. 'bat wings']

(32) *oko   na-kɨ*
      crab   tooth-PL
      'pliers' [lit. 'crab teeth']

(33) *yõra-kɨ-rimë   a*
      antler-PL-ATTRIB SG
      'pick-ax' [lit. 'an antlers-like thing']

(34) *tëpë   hwasipë*
      anteater   back
      'a type of concave hoe' [lit. 'anteater's back']

(35) *tẽremë   nahasi*
      mythological.being claw
      'a type of triangular hoe' [lit. 'claw of a specific mythological being']

(36) *xama moka*
      tapir penis
      'a type of hoe for digging' [lit. 'tapir's penis']

Examples (37) and (38) show the adoption of the Portuguese words *carro* 'car' and *camisa* 'shirt', respectively, while at the same time incorporating the new concepts through the use of body-part metaphors.

(37) *kahu*   *mahɨ=kɨ*
       car (Port. *carro*) foot-PL
       'car tires' [lit. 'car's feet']

(38) *kapixa*   *mamo-kɨ*
       shirt (Port. *camisa*) eye-PL
       'shirt buttons' [lit. 'shirt's eyes']

6.   Euphemisms for death

Throughout the previous sections, examples of new vocabulary as well as established Yanomae expressions have been presented that support Deutscher's claim that "…if not from the physical world, where else could terms for abstract concepts come from?… The mind cannot just manufacture words for abstract concepts out of thin air – all it can do is adapt what is already available" (2005: 127). The examples in this section focus on the most abstract of concepts that confront human societies: death. This highly sensitive abstraction is communicated in the Yanomae language through an extension of the concept of metaphor into the related phenomena of euphemism and metonymy. Allan and Burridge define euphemism as "an alternative to a dispreferred expression, in order to avoid loss of face" (1991: 11). Burridge further clarifies the term "dispreferred expression" by describing it as "one that is not desired or appropriate on a given occasion. Typically it denotes a taboo topic and so might alternatively be called a 'taboo term'" (2006a: 455). The original Tongan conception of *tabu* as "forbidden behavior, in particular, behavior believed to be dangerous to certain individuals or to the society as a whole" (ibid.) is closer to the Yanomami severe prohibition of speaking of the dead than the common use of euphemism in contemporary contexts, such as English *sanitation worker* for *garbage man* or *dearly departed loved one* instead of *dead relative*.

Among the Yanomami of Amazonia, addressing someone in public by his or her personal name is an insult; likewise, mentioning the name of a deceased person is an offense akin to blasphemy in Western cultures. Burridge (2006b: 452) notes the close connection between naming taboo, the topic of death, and a fear of evil or supernatural powers in many cultures, and Yanomae figurative language nicely illustrates this intersection between linguistic usage and cultural beliefs. Unlike many languages that have euphemisms for death and dying that offer an

image of consolation by emphasizing "different aspects of the physical event …
for example, death as a journey (*pass away*) or death as the beginning of a new life
(*go to a better place*)" (Burridge 2006a: 459), death in the Yanomae culture is the
end of a productive life.

The abstract concept of *death* is "repackaged" in concrete terms to make it
more acceptable in normal conversation. In Yanomae, the domain of *death* is
"metaphorically linked with" (Coulson 2006: 33) the domain of *daily activities*,
which are further specified according to age and gender roles and identified with
particular objects – in this case, specific baskets and arrows. Dying is an aber-
ration of daily activities, and, thus, it is euphemistically referred to within that
"normal" context, instead of directly using Yanomae lexemes that refer to dying,
death, or a dead human body, which would be culturally taboo.

Death may be conveyed as emptiness, loss, or deterioration in the domestic
space. An empty hearth, as in example (39), with no fire or food cooking is not
normal in daily life and serves as an indication of the permanent absence of a
person who otherwise would inhabit the space. In examples (40) and (41) dete-
rioration is the metonymic image that expresses a death, as the bark (hammock)
and old house post represent the dead people. In addition, an understanding of
cultural assumptions within these examples suggests that the dead person in each
case was elderly. The use of the term *rainathe* in (40) suggests a traditional bark
hammock, which would only be used by an older person; cotton or manufactured
hammocks, *toutou siki*, are much more common these days. In (41) the fact that
the house post that fell is described as old, *pata*, implies that the dead person is an
elderly man, keeping in mind that houses are constructed by men.

(39) *yutu   nahi   proke*
      long.ago hearth empty
      'For a long time [her/his place at] the hearth [has been] empty.'

(40) *rainathe   prohe-pra-rio-ma*
      bark   be.loose-INTEN-TEL-PAST
      'A (traditional) bark (hammock) became very loose.'

(41) *pata nahi   ke-rayo-ma*
      old   house.post fall-TEL-PAST
      'The old house post fell.'

The metonymic images found in Yanomae euphemisms recall 16th century Na-
huatl use of "source material from the physical world, scenes of life, social roles…"
to create a complex system of metaphors rooted in a common cultural context
(Palmer 1996: 240). An examination of the use of euphemism in the Yanomae
language demonstrates how *not* speaking of the dead can provide insights into

Table 1. Euphemistic metonymy

	– ADULT	+ ADULT	
physical object dead person	small arrows *ruhu masi* (boy)	long arrows *xarakaki* (man)	+ MALE
physical object dead person	small basket *xote he* (girl)	large basket *wii a* (woman)	+ FEMALE

important aspects of Yanomami culture. Euphemism can be expressed through metonymy as "the conceptual principle in which the object replaces the user" (Shen 2006: 462). Table 1 schematizes information about the gender and the relative age (adult vs. child) of the deceased that is conveyed by reference to specific items of material culture (baskets and arrows) that identify the individual's role in traditional Yanomami society. In metonymy "[s]ome subpart of a thing or aspect of a relation comes to stand for the whole" (Rice 2012: 25). As a classic example of metonymy, the person, in death, "becomes" the object with which it is identified in life.

Several verbs may be used euphemistically to announce a death when combined in an utterance with one of the following four specific objects: *wii a* 'a large tightly-woven carrying basket' used by adult women to bring produce from the garden, *xote he* 'a small shallow basket' used by young girls, *xarakaki* 'long (cane) arrows' used by adult men for hunting game, and *ruhu masi* 'a small (palm) arrow' used by young boys learning to hunt (i.e. small birds, insects, or lizards). These four objects are commonly used in daily, village life and are associated with activities of specific gender and age groups. In example (42) the death of a young boy is announced in terms of the corresponding metonym, a small arrow:

(42) *ruhu=yama=masi kasi-ma-re-ma*
       arrow=we=CL   edge-CAU-TEL-PAST
       'We put small arrows (= a boy's body) at the edge (of the village).'

In the context of euphemistic usage, the verbs *kasima-* 'to put at the edge (outside the village),' *yokama-* 'to put aside, to put to the side' and *urihima-* 'to put in the forest' refer to the initial stage in the Yanomami funerary ritual whereby the cadaver, carefully wrapped in a bundle of leaves, sticks, and vines, is placed outside the village in the forest to await decomposition (Albert & Gomez 1997: 169).

Similarly, the four metonymic domestic objects may be used with other verbs, such as 'carry' *hiipu-* in (43) and 'place' *itha-* in (44) to announce the deaths of a young boy or an adult woman, respectively, using different contextual imagery. Example (43) includes an additional metaphor whereby the physical act of *carrying*

is extended to mean 'keep in one's heart or memory.' This same metaphorical use of the verb *carry* is common in English.

(43)   *kama=hwɨ̃=e-nɨ*   *ruhu=masi*   *hĩĩpu xoa*
       3SG=father=POSS-ERG small arrow=CL.plant carry still
       'His father still carries the small arrows.'

(44)   *okapɔ-ha-ixoru-nɨ*   *wɨɨ=yama=a ĩtha-ke-ma*
       sorcerers-SEQ-attack-SEQ basket=we=SG place-FOC-PAST
       'After the sorcerers attacked, we placed (firmly on the ground) a large basket
       (= a woman's body).'

The use of euphemism and metonymy in Yanomae as a solution to linguistic taboos dealing with death poignantly illustrates the interwoven relationship between language and culture. The discarded objects (baskets and arrows) of everyday use become, through language, symbols of the dead children and adults. In example (45) the small shallow basket *xotehe* is a young girl. Moreover, additional clues to the identity of the dead person may be revealed by a deeper knowledge of the culture that accompanies the metonymic expressions, as already mentioned in the discussion of examples (40) and (41). In example (46) the specific location *xoa e a hēhāowiha* 'from the space next to my father-in-law's' suggests that the dead woman was probably the speaker's wife or sister-in-law, given the traditional matrilocal residence pattern of Yanomami groups.

(45)   *xote=yama=he*   *yoka-ma-re-ma*
       small.basket=we=CL put.aside-CAU-TEL-PAST
       'We put aside a small basket (= a girl's body).'

(46)   *xoa=e=a*   *hēhāowi=ha*   *wɨɨ=yama=a*
       father-in-law=POSS=SG interior.living.space=LOC large.basket=we=SG
       urihi-ma-re-ma
       forest-CAU-TEL-PAST
       'From the space next to my father-in-law's, we put a large basket (= a woman's body) in the forest.'

The urgency of the directive expressed by the use of the imperative in example (47) conveys the importance of the task, yet the literal meaning of the words would seem nonsensical if the hearer were unaware of the euphemistic metonym. Likewise, the meaning of example (48) would not be obvious without an understanding of the euphemistic use of the verb *hatëtë* 'clasp tightly' in the context of the metonymic 'small arrows.' The listener would not realize that the speaker, who is not physically holding small arrows, was actually saying, 'Now my son has died (or my son's funeral has been planned).' All this, of course, is perfectly clear to a native speaker of Yanomae.

(47) *rope   ruhu=wama=masi   urihi=a   ha-i-ma-ri*
      quickly arrow=you.PL=CL.plant forest=SG go.out-DYN-CAU-IMP
      'Quickly, take the small arrows (= the boy's body) to the forest!'

(48) *hwei=tëhë ipa ruhu=ya=masi   hatëtë-ma-ke-ma*
      this=TEMP my small.arrow=I=CL.plant clasp.tightly-CAU-FOC-PAST
      'Now I kept tightly clasped my small arrows.'

7. Concluding thoughts on endangered metaphors

The purpose of this paper is to illustrate the use of metaphor as it is commonly used in a lesser known language of Amazonia and to show how the language and culture of the Yanomae are tied to the rain forest environment not only as reflected in specific lexical domains, but also as seen through the use of euphemistic metonymy to overcome a taboo on death by *not* speaking of the dead. It is clear in the use of body-part metaphors to refer to traditional objects, geographical phenomena, and neologisms that the bodies of humans and tropical fauna serve as reference points for Yanomae speakers. Furthermore, orientational and temporal metaphors reflect a vibrant traditional culture. This culture and, consequently, the language through which it is expressed will survive and have meaning only as long as its rain forest habitat remains intact. Andreas Musolff proposes "a system of criteria for classifying metaphors which can be said to be 'endangered' in a linguistic-ecological sense" (Abstract LAUD 2014: 109). The idea of linguistic-ecological endangerment is particularly relevant to languages in the Amazon Basin of South America, which are seriously threatened not only by encroaching national languages (especially Portuguese and Spanish) but also by the invasion and seizure of their lands by settlers, ranchers, loggers, and wildcat gold miners as well as international corporations seeking to exploit natural resources, especially oil.

Jonathan Loh and David Harmon emphasize that the decline in biodiversity globally coincides with the loss of linguistic diversity, especially in the Americas and Australia, where they claim "the most highly threatened language families are" (Abstract LAUD 2014: 93). The Yanomami language family is among these, as its population of speakers continues to suffer from uncontrolled invasions by outsiders into their territory. Unwarranted contact with outsiders not only results in changes to the traditional lifestyles, but it also brings disease and death to remote Yanomami villagers, who do not have acquired resistance to common Western infectious diseases. As one of the largest of the least acculturated indigenous groups in the Amazon today, the approximately 33,000 Yanomami in Brazil and Venezuela are especially endangered physically and culturally as their mineral rich lands are coveted by individuals as well as their respective governments. This

paper provided a short description of some of the metaphors in one of the four main Yanomami languages; it does not pretend to even begin to show the richness of the linguistic structures and the knowledge that the speakers of these languages have about the flora and fauna of the rain forest they inhabit. It remains to be seen whether the non-indigenous, so-called developed, world will learn to value such knowledge and preserve the world's remaining biolinguistic and human diversity.

Abbreviations

CAU	causative	POSS	possessive
CL	classifier	SEQ	sequential
DYN	dynamic	SG	singular
ERG	ergative	STATIVE	stative
FOC	focalizer	TEL	telic
IMP	imperative	TEMP	temporal
INTEN	intensifier	1	1st person
PAST	past	3	3rd person
PL	plural		

Acknowledgements

References

Abstracts LAUD 2014. 36th International LAUD Symposium. March 31–April 3, 2014. Landau: University of Koblenz-Landau.

Albert, B. & Gómez, G.G. 1997. *Saúde Yanomami: Um manual etnolingüístico. Coleção Eduardo Galvão*. Belem: Museu Goeldi/Paris: ORSTOM.

Allan, K. & Burridge, K. 1991. *Euphemism and Dysphemism: Language Used as Shield and Weapon*. Oxford: OUP.

Burridge, K. 2006a. Taboo euphemism, and political correctness. In *Encyclopedia of Language and Linguistics*, Vol. 12, K. Brown (ed.), 455–462. Oxford: Elsevier. doi:10.1016/B0-08-044854-2/01092-0

Burridge, K. 2006b. Taboo words. In *Encyclopedia of Language and Linguistics*, K. Brown (ed.), 452–455. Oxford: Elsevier. doi:10.1016/B0-08-044854-2/00778-1

Coulson, S. 2006. Metaphor and conceptual blending. In *Encyclopedia of Language and Linguistics*, Vol. 8, K. Brown (ed.), 32–462. Oxford: Elsevier. doi:10.1016/B0-08-044854-2/01098-1

Deutscher, G. 2005. *The Unfolding of Language: An Evolutionary Tour of Mankind's Greatest Invention*. New York NY: Owl Books/Henry Holt & Company.

Gomez, G.G. 1990. The Shiriana Dialect of Yanam (Northern Brazil). PhD dissertation, Columbia University.

Gómez, G.G. & Albert, B. Léxico themático da língua Yanomae. Ms.

Lakoff, G. & Johnson, M. 1980. *Metaphors We Live By*. Chicago IL: University of Chicago Press.

Loh, J. & Harmon, D. 2014. Comparing status and trends in linguistic and biological diversity. *Abstracts*, 91–93.

Musolff, A. 2014. In which sense(s) can we speak of 'endangered metaphors'? *Abstracts*, 108–110.

Nkamigbo, L.C. 2014. Metaphorical proverbs in Koring. *Abstracts*, 113–114.

Palmer, G. 1996. *Toward a Theory of Cultural Linguistics*. Austin TX: University of Texas Press.

Rice, S. 2012. Our language is very literal: Figurative expression in Dene Sųłiné [Athapaskan]. In *Endangered Metaphors* [Cognitive Linguistics Studies in Cultural Contexts 2], A. Idström & E. Piirainen (eds.) in cooperation with T.F.M. Falzett, 21–76. Amsterdam: John Benjamins. doi:10.1075/clscc.2.03ric

Shen, Y. 2006. Figures of speech. In *Encyclopedia of Language and Linguistics*, Vol. 4, K. Brown (ed.), 459–464. Oxford: Elsevier. doi:10.1016/B0-08-044854-2/00510-1

Measuring and understanding ethnolinguistic vitality in Papapana*

Ellen Smith
University of Newcastle

This paper presents a sociolinguistic profile of Papapana, spoken in Bougainville, Papua New Guinea, and investigates why and to what extent Papapana is endangered. Language endangerment is the result of complex and interrelated sociolinguistic variables. The study of language contact and use in the Papapana speech community demonstrates the unique ways in which these variables interact and provides crucial insights into factors affecting linguistic vitality in general. The assessment of Papapana using some of the many ethnolinguistic vitality assessment frameworks challenges the assumptions and definitions of these frameworks, and suggests they need re-evaluating, particularly as they often fail to distinguish symptoms and causes of linguistic endangerment. It is vital to make this separation as predictive and diagnostic assessments may have different applications.

Keywords: causes, endangerment, ethnolinguistic vitality, language contact, Papua New Guinea, status, use

1. Language endangerment and death

Over the past few decades, interest in ethnolinguistic vitality has grown due to the realisation that many of the world's 6,000 to 7,000 languages are endangered. Estimates range from 50% loss (Crystal 2000: 19) to 90% loss (Krauss 1992: 7) of the world's languages in the coming century. Languages are thus disappearing at an alarming rate: according to these estimates, between two and four languages will die every month for the next 100 years.

The realisation of the extent of global language endangerment has led to increased appreciation for what is lost when a language dies. Linguistic scholarship depends on linguistic diversity to gain a fuller picture of the human capacity for language, to develop grammatical theories and to classify languages, but it is of course not only linguists who are disadvantaged when a language disappears. The demise of a language may result in the loss of scientific, cultural or historical information encoded in the language itself, while the transmission of information, memories and stories may not be possible if the language used to impart this knowledge is no longer understandable. Furthermore, since "to choose to use a language, is an act of identity" (Nettle & Romaine 2000: 173), a speech community might experience a loss of ethnic and cultural identity if their language dies.

The study of language death and endangerment is crucial if one wishes to document, maintain or revitalise a language. Investigating the causes of linguistic endangerment is necessary for preventing or reversing the demise of a language, while assessing the extent to which a language is endangered allows us to identify and prioritise the languages that are most in need of documentation and/or revitalisation. Since the process of language death is complex and varied, the study of language use in individual speech communities may provide crucial insights into factors affecting linguistic vitality. Indeed, Foley (2004: 30) argues that we need more "studies of speech communities in transition to understand the hows and whys of language endangerment".

This paper addresses Foley's call by presenting a sociolinguistic profile of the endangered language Papapana (Northwest Solomonic, Oceanic, Austronesian) spoken on the northeast coast of Bougainville island, Papua New Guinea (PNG). Using data from a doctoral documentation and description fieldwork project, this paper investigates why and to what extent Papapana is endangered.

2.   Papapana sociolinguistic milieu

This paper is based upon data collected during two fieldwork trips: June 2011 to March 2012, and March to May 2013. While the project focused on documenting Papapana to core documentation level, writing a grammatical description and creating community materials, I also compiled sociolinguistic profiles of the Papapana community members and documented Papapana's socio-cultural context. This data was obtained through participant observation, informal interviews, and compiling genealogical and sociolinguistic profiles of around 800 individuals who lived in the Papapana villages or were closely related to its inhabitants. I was not able to meet every single individual so some of the information was obtained from their relatives, and while I tried to be as accurate as possible, it should be noted

that some of the data is approximate. For speakers who participated in audio re-
cordings, their competency in Papapana was self-evident but for others I relied on
speakers' judgements of their own competency or on their family's judgements.
This section reports the findings on speaker location and numbers (§2.1), lan-
guage contact (§2.2) and language use (§2.3), in order to provide the information
necessary for investigating the causes and extent of Papapana's endangerment.

2.1   Speaker location and numbers

Papapana is spoken on the northeast coast of Bougainville island, the Autono-
mous Region of Bougainville, PNG (see Map 1), an area of the Pacific known as
Melanesia. The Papapana speech community originates in Teperoi village but is
currently also located in five other villages nearby: Peuni, Koikoi, Maras, Barora
and Iraka (see Map 2 and Map 3). Teperoi is situated along a 1.5 kilometre track.
Peuni, Barora and Iraka are each situated in one clearing, while Koikoi and Maras
consist of six and seven sites respectively.

In May 2013, the total number of fluent, first language (L1) Papapana speak-
ers with full productive ability was 106, there were fifty-five second language (L2)
or semi-speakers with partial productive ability, and there were around 136 peo-
ple with only passive understanding of Papapana. Table 1 shows the number of
these different types of speaker by location. If one considers only the speakers
who were residing in the six Papapana villages, L1/fluent Papapana speakers ac-
counted for 17%, L2/semi-speakers accounted for 8%, and people with passive

Map 1. Papapana in Papua New Guinea

Map 2. Papapana in Bougainville

knowledge of Papapana accounted for 21% of the population. If one includes the speakers who were living elsewhere (since they may visit the Papapana villages), fluent Papapana speakers constituted 21%, semi-speakers 11% and people with passive knowledge 27% of the total population.
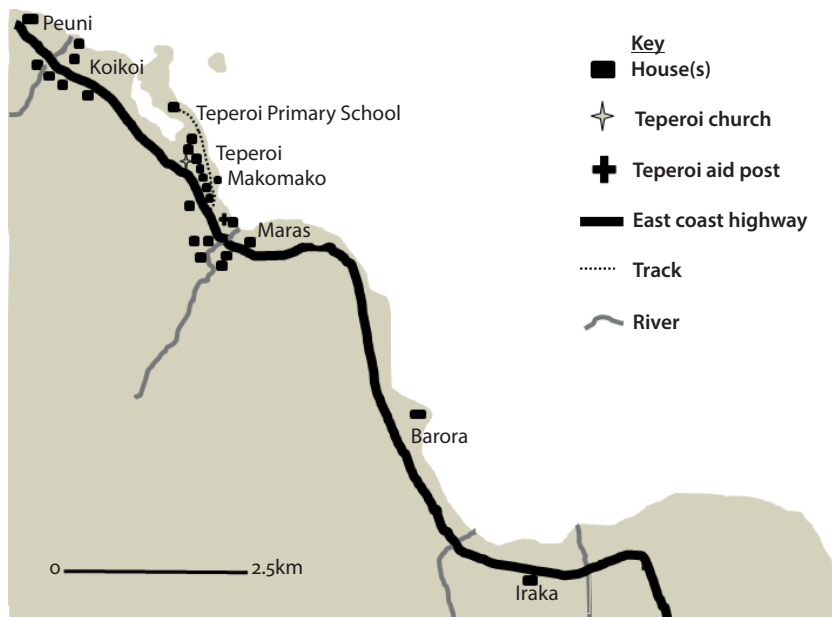
Map 3. Papapana villages

Table 1. Speakers by location

Location	Population of location	L1 speakers		L2/semi-speakers		Passive	
Peuni	21	3	14%	4	19%	7	33%
Koikoi	47	7	15%	3	6%	10	21%
Teperoi	224	47	21%	17	8%	45	20%
Maras	91	13	14%	7	8%	14	15%
Barora	68	10	15%	9	13%	24	35%
Iraka	59	7	12%	2	3%	7	12%
Elsewhere		19		13		29	
TOTAL	510	106	17–21%	55	8–11%	136	21–27%

2.2   Language contact

This section describes the languages of PNG and Bougainville (§2.2.1), and the history of language contact in the Papapana community, from pre-colonisation (§2.2.2), European colonisation (§2.2.3), national independence (§2.2.4) and the Bougainville civil war (§2.2.5), to the present day (§2.2.6).

2.2.1   *Linguistic diversity*

PNG is one of the most linguistically diverse countries in the world: 836 languages are spoken by six to seven million people. Around 230 languages belong to the Oceanic subgroup of the Austronesian family, while the remaining 600 or so languages are non-Austronesian (Wurm 2003: 25), or *Papuan*. The official languages of PNG are English, the creole Tok Pisin and the pidgin Hiri Motu. Aside from these, Papua New Guinean languages generally have very small speaker bases: more than half have less than 1000 speakers (see Wurm 2003: 25).

The Autonomous Region of Bougainville has a population of 234,280 people (2011 census) and is home to twenty-three local languages: eight Papuan and sixteen Austronesian. Table 2 shows each language's speaker numbers (Lewis, Simons & Fennig 2014) and subgroup, while Map 4 shows the location of the languages spoken on Buka and Bougainville islands. The Ellicean languages and Nehan are spoken on atolls. Papapana is therefore clearly situated in a *residual zone*, that is,

Table 2. Bougainville languages speaker numbers

Language group		Name	Speakers
Papuan (69,000 speakers)	South Bougainville	Buin/Telei/Kugara	26,500
		Nasioi/Kieta	20,000
		Motuna/Siwai	6,600
		Nagovisi/Sibbe	6,000
	North Bougainville	Rotokas	4,320
		Konua/Rapoisi	3,500
		Eivo/Askopan	1,200
		Keriaka/Ramopa	1,000
Austronesian (53,556 speakers)	Northwest Solomonic (Western Oceanic)	Halia	25,000
		Nehan/Nissan	6,500
		Teop	5,000
		Tinputz/Vasui	3,900
		Solos	3,200
		Petats/Majugan	2,000
		Saposa	1,400
		Hahon	1,300
		Banoni/Tsunari	1,000
		Torau/Rorovana	600
		Piva/Lawunuia	550
		Papapana	106
		Uruava	0
	Ellicean (Central-Eastern Oceanic)	Takuu/Mortlock	1,750
		Nukumanu/Tasman	700
		Nukuria/Nahoa	550

Map 4. Oceanic and Papuan Languages of Bougainville and Northwestern Solomon Islands (FROM Evans 2009)

a zone which has "high diversity… [is] inhabited by small groups… with many different language types, among whom… multilingualism is the norm" (Nettle & Romaine 2000: 38).

2.2.2   *Pre-colonisation*

Buka and Bougainville islands were first populated by Papuan language speakers around 30,000 years ago (Tryon 2005: 33). Around 3,000 years ago, Austronesian language speakers arrived from the north (Regan & Griffin 2005: 475) and continued south into the Solomon Islands and beyond. Much later, the descendants of some of those who had settled on the islands immediately south of Bougainville, resettled along Bougainville's eastern coast (Oliver 1991: 3). Papapana ancestors were part of this migration from the south in the mid-19th century. Prior to European settlement, Papapana speakers only had contact with Rotokas speakers for the purposes of trade and intermarriage. Rotokas was the intergroup language since Rotokas speakers greatly outnumbered Papapana speakers.

2.2.3   *European colonisation*

With European colonisation came the introduction of a plantation economy. The Numa Numa plantation had been established north of Teperoi by 1912 (Laracy 2005a: 140) and Koikoi, Tenekau, Kurwina and Arigua plantations north and south of Teperoi by 1968 (see Map 2). The manual labour in these plantations was mainly supplied by indigenes (Oliver 1991: 31). Papapana speakers reported that many Buin and Motuna (South Bougainville, Papuan) speakers were recruited to work in the plantations surrounding Teperoi. Due to their proximity to the plantations and/or being labourers themselves, Papapana speakers experienced increased contact with many other speech communities, and there was increased exposure to Tok Pisin, whose origins largely lie in Samoan plantation activities in the 1880s (Wurm 1979: 6).[1]

European settlers also introduced their religious beliefs and established churches and schools at mission stations. The Catholic Society of Mary (Marists) established the first mission station near Kieta in 1901 (Laracy 2005b: 126), while the Methodists arrived in 1922 (Laracy 2005b: 126) and the Seventh Day Adventists in 1924 (Regan & Griffin 2005: 476). The closest Marist missions to Teperoi were Asitavi (established 1935) and Mabiri (established 1958) (Laracy 1976) (see Map 2) but the Papapana people had their own Catechists in the village by 1931 (McHardy 1935). At least until 1914 Catholic missionaries in PNG used their parishioners' languages, whereas Protestant missionaries resorted to a lingua franca (Ross 1996: 595). Nevertheless, the arrival of Protestant competition in Bougainville prompted Marists to make greater use of Tok Pisin and teach English in some mission schools (Laracy 2005b: 126). By 1968 the Bible Society of PNG had translated the New Testament into Tok Pisin, and the whole Bible by 1989.

---

1.   For more information on the origins of Tok Pisin, see Baker and Mühlhäusler (1996), Mühlhäusler (1976) and Wurm (2007: 444).

2.2.4   *Educational policies and national independence*

English teaching and the use of English as a medium of instruction were the official educational policy from 1956 (Ross 1996: 597). After national independence in 1975, PNG educational policies focused on raising literacy levels in some local languages, Tok Pisin, Hiri Motu and English; however, hundreds of other local languages still remained unrepresented in elementary education (Wurm 2007: 445). A 1997 declaration, by PNG's Vice-Minister of Education Dr Jon Waiko, made each of PNG's local languages official languages that could be used in basic elementary education, along with Tok Pisin, Hiri Motu and English (Wurm 2007: 445).

2.2.5   *Panguna mine and the Bougainville Crisis*

In 1969 Bougainville Copper Limited (BCL) established a mine in Panguna, Southern Bougainville (see Map 2). When the Panguna mine began production in 1972, it was the world's largest open-cut copper mine and generated nearly half of PNG's export revenue. The mine increased the use of Tok Pisin in the region: "in the absence of [a dominant local] language, it was Pidgin which became the common means of communication with the majority of employees and… villagers" (Vernon 2005: 263).

Panguna mine became politically contentious with disputes over land tenure and allegations of environmental damage and inequitable distribution of mining revenues. This prompted a civil war, the Bougainville Crisis, which lasted from 1989 until 1997. In total, 10,000 to 15,000 people died. The Crisis has caused huge social and economic change, massive population displacement, a breakdown in law and order, and a decline in the educational system. Displacement increased contact with other local languages and increased the need for Tok Pisin as a lingua franca, especially in the PNG government Care Centres.

2.2.6   *The 21st century*

Papapana speakers once occupied a larger area of coastal land, from Kiviri in the north to Arigua plantation in the south (see Map 2). The establishment of plantations, and government services such as Wakunai airstrip, decreased the land size occupied by Papapana speakers. The knock-on effect was that by the mid-20th century, Teperoi was overpopulated. Certain families decided to reclaim their ancestors' land and settle the villages of Maras, Barora and Iraka in the mid-20th century, and Peuni and Koikoi from 1990 to 2010. This dispersal has increased contact with other speech communities.

Since the Bougainville Crisis, there has been an increase in permanent movement into and out of the Papapana community for intermarriage among people from diverse linguistic backgrounds. In May 2013, 30% of fluent Papapana speakers were married to each other in intra-ethnic marriages, 5% were Papapana-

Table 3. Multilingualism in Papapana community

L1	Speakers	Proportion
Tok Pisin	338	66%
Papapana	87	17%
Rotokas	20[6]	4%
Halia	13	3%
Motuna	12	2%
Buin	8[3]	1.5%
Torau	6[5]	1.1%
Nasioi	5[1]	1%
Nehan	5	1%
Other (PNG)	5	1%
Tinputz	3[2]	0.6%
Nagovisi	3	0.6%
Other (Solomon Islands)	2	0.4%
Teop	1[1]	0.2%
Eivo	1	0.2%
Banoni	1	0.2%
TOTAL	510	

speaking widows who had been married to Papapana-speaking men and 56% of fluent Papapana speakers were married in inter-ethnic marriages to speakers of over ten different local languages. Some people had also moved into the Papapana community for work and consequently there were over fourteen local languages represented. Table 3 shows the number of speakers of different L1s who lived in the six Papapana villages. Where speakers were multilingual, they are included in the main count for their primary language and in square brackets for the language they identified with least. The numbers in square brackets are not considered in the calculation of the proportion of speakers out of the total population of 510 as otherwise there would be more than 510 tokens.

There has also been increased temporary or permanent movement of Papapana speakers to population centres for employment, markets, educational and medical services, entertainment, or important religious celebrations in larger parishes. It has become easier to travel to towns such as Buka and Arawa, government stations such as Wakunai, and mission compounds and local high schools such as those at Asitavi, due to Public Motor Vehicles (PMVs) travelling daily between Arawa and Buka, local PMVs travelling around the Wakunai district, and the completion in 2012 of a bridge network along the east coast highway. This mobility has increased contact with other speech communities.

Media and technology have also increased contact. In PNG, television is almost exclusively in English, radio is in English, Tok Pisin and Hiri Motu or, for provincial radio, the larger local languages, while newspapers are in English or Tok Pisin with some provincial newspapers in a local language (Lynch 1998: 268). I never saw a newspaper or radio in the Papapana villages, but some community members have mobile phones and in 2013, several teachers had DVD televisions which they would use (power permitting) to watch films or music videos in Tok Pisin, English or other local languages.

2.3 Language use

This section describes Papapana language use between 2011 and 2013 in the following domains: the home (including intergenerational transmission) (§2.3.1), education (§2.3.2), work (§2.3.3), administration (§2.3.4), religion (§2.3.5), and sports events and media (§2.3.6).

2.3.1 *Home and intergenerational transmission*
In the Papapana villages, being *at home* means being around the buildings in which people sleep and cook, and having contact with grandparents, parents, children, siblings, cousins, in-laws etc. Papapana is used among fluent and semi-speakers in the home domain but Tok Pisin is used when non-Papapana speakers are present.

The language spoken in the home domain relates to intergenerational language transmission, that is, "whether parents and older members of the community are speaking the language with and around children and young people" (Florey 2005: 45). Taking into account all the individuals who know Papapana to some extent, Table 4 presents the distribution of L1/fluent speakers, L2/semi-speakers and

Table 4. Speakers by age

Age group		L1 speakers	L2/semi-speakers	Passive	TOTAL
Children	0–9	2	1	7	10
	10–19	0	7	48	55
Parents	20–29	3	13	45	61
	30–39	23	33	33	89
Grandparents	40–49	40	1	2	43
	50–59	24	0	0	24
Great-grandparents	60–69	11	0	1	12
	70–79	1	0	0	1
	80–89	2	0	0	2
TOTAL		106	55	136	297

Table 5. Proportion of speakers in age group

Age group		Total population of age group in Papapana villages	L1 speakers		L2/semi-speakers		Passive	
Children	0–9	114	2	1.8%	0	–	7	6%
	10–19	144	0	–	8	6%	46	32%
Parents	20–29	76	3	4%	9	12%	33	43%
	30–39	81	17	21%	24	30%	18	22%
Grandparents	40–49	49	33	67%	1	2%	2	4%
	50–59	31	20	65%	0	–	0	–
Great-grandparents	60–69	12	9	75%	0	–	1	8%
	70–79	1	1	100%	0	–	0	–
	80–89	2	2	100%	0	–	0	–
TOTAL		510	87		42		107	

people with passive knowledge of Papapana by each age group: 95% of L1/fluent speakers are above the age of thirty, while most L2/semi-speakers and people with passive knowledge are under forty.

Taking into account all individuals in the six Papapana villages, Table 5 presents the proportion, within each age group, of the Papapana speakers who live in those villages. The proportion of L1/fluent speakers per age group significantly increases with increased age. The proportion of parents and children who are L2/semi-speakers is higher than those who speak it fluently, while hardly anyone over forty speaks Papapana partially or has passive knowledge, suggesting that it is all or nothing for older speakers.

In 2013 there were some speakers in the parental generation who had had Papapana transmitted to them but, with the exception of one couple (who had two Papapana-speaking sons under ten), were not transmitting the language to their own children. Consequently, grandparent and great-grandparent generations are generally the fluent speakers, parents in their 30s are often semi-speakers, while younger parents and children have passive knowledge of Papapana at best.

2.3.2 *Education*

In 2013, Teperoi Primary School had been following the 2003 PNG National Department of Education Elementary Language Syllabus which states that "The students' first language is to be used as a medium of instruction for the first three years of education" (NDOE 2003: iv). The students' L1 is referred to as their "vernacular", which is defined as the "tokples" (NDOE 2003: 1). In Teperoi, Papapana is the "tokples" and it is difficult to achieve the aforementioned objective as

the students have different linguistic backgrounds (as do the teachers) and many speak Tok Pisin as their L1. In 2011 both Elementary Prep and Elementary 1 were taught by fluent Papapana speakers, but in 2012 and 2013, only Elementary Prep was. In Elementary Prep language classes, Papapana and English are taught as L2s, with the medium of instruction being Papapana and Tok Pisin. Resources in Papapana, such as story books or paper flashcards, are created by the teacher and are of a temporary nature. Some songs and games are conducted in Papapana. In other classes, Papapana is sometimes used for numbers and naming objects, but Tok Pisin is the medium of instruction.

2.3.3   *Work*
Within the villages, work includes housework, building shelters, fishing, hunting, gardening and copra/cocoa production. The latter four may also take place outside of the village. Another occupation within the villages is teaching, while some community members work in other towns as tradesmen, teachers, nurses or in local government. Work inside the village could be considered part of the home domain. In the external work domain, Papapana might be used among fluent speakers but it is not the language of that domain (which may be Tok Pisin, English or another local language) and Papapana is less likely to be used given the increased chance of non-Papapana speakers being present.

2.3.4   *Administration*
In May 2013, all the main chiefs lived in Teperoi or Maras, while each village had a clan chief. All chiefs are fluent Papapana speakers and meet once a month. Without being privy to these meetings, I cannot say what language is used. Speakers said they use Papapana, but Tok Pisin may well be used too. Monthly meetings and weekly announcements concerning the entire community are conducted in Tok Pisin. Regional government and administration outside of the six Papapana villages is most certainly conducted in Tok Pisin, or another local language.

2.3.5   *Religion*
Teperoi village has a Catholic church. Church sermons, readings and prayers are conducted in Tok Pisin. Songs are conducted in Tok Pisin, and occasionally in English or Papapana. According to one speaker, traditional ceremonies outside church are conducted in Papapana.

2.3.6   *Sports events and media*
On most Sunday afternoons, community members play sports on the school grounds. Community members also meet for fundraising events. Social events such as these which take place within a Papapana village are conducted in a

mixture of Papapana and Tok Pisin depending on a particular conversation's participants. Social gatherings which take place outside of the six villages are conducted in Tok Pisin, though Papapana speakers might speak Papapana among themselves. Papapana is not represented in any media.

3. Investigating causes of endangerment

Wurm (2007: 530) describes Papapana as being under pressure from Tok Pisin and large related NWS languages such as Tinputz. This section assesses the accuracy of this description by firstly considering the types of and motivations for language shift (§3.1) before investigating causes of language shift in the Papapana community (§3.2).

3.1 Language shift: Types and motivations

Except for the rare situation in which a speech community is physically eliminated, all instances of language death are due to language shift. Language shift is defined as "a change in the balance of domain-particular use of languages in the multilingual group's repertoire" (Matras 2005: 238) and this often results in "partial or total abandonment of a group's native language in favour of another" (Winford 2003: 15). A prerequisite for language shift is language contact; shift from one language to another can only occur if there is exposure to other languages.

There are two types of language shift defined in the literature; forced and voluntary (Nettle & Romaine 2000: 90–91; Campbell & Muntzel 1989: 183–186), though in reality this distinction might not be as clear-cut. In forced shift situations, a dominant language group may make their language compulsory and punish the use of the minority group's language. To avoid persecution, minority language speakers cease using their language. Today however, language shift is most often voluntary: "a community… perceive that they would be better off speaking a language other than their original one" (Nettle & Romaine 2000: 91).

There are two subtypes of voluntary shift, bottom-up and top-down (Nettle & Romaine 2000). In bottom-up shift, the language is lost in most everyday domains but survives in ceremonial or formal domains. In top-down shift, the language disappears first from official institutions and public domains (in which there is shift to a language of wider communication) and is eventually restricted to the home domain.

In cases of voluntary language shift, the biggest question is why a community perceives that they would be better off speaking a language other than their original one. Patterns of language choice reflect language attitudes thus "shift in language is

caused… by shifts in personal and group values and goals" (Kulick 1992:9). Such attitudes pertain to the usefulness and worth of the language: "speakers abandon their native tongue in adaptation to an environment where use of that language is no longer advantageous" (Grenoble & Whaley 1998:22). This is widely accepted as the fundamental cause of voluntary language shift; however, the complicated issue is identifying the environmental changes that bring about decreased efficacy of a language (Grenoble & Whaley 1998:22). Environmental changes may include changes in the speech community's demographic composition, culture or economic base, or changes in institutional policies. Identifying these changes is challenging because it involves a complex constellation of interrelated sociolinguistic variables that are specific to a particular language contact setting.

3.2   Language shift in the Papapana community

Until the 1990s, PNG had the lowest level of language endangerment of all areas in the world containing many small languages (Wurm 2007). Nevertheless, by the beginning of the twenty-first century, language endangerment had increased and today in PNG, sixteen languages are extinct, seventy-seven languages are endangered and 200 are likely to become endangered (Wurm 2007:445). The threat is the increased use of Tok Pisin: Tok Pisin is now spoken by more than three-quarters of the population, compared with half the population twenty years ago (Wurm 2007:444) and an increasing number of younger Melanesians grow up speaking Tok Pisin exclusively, or with greater confidence than their parents' vernaculars (Lynch, Ross & Crowley 2002:28).

In the Papapana community, there has been significant voluntary top-down shift from Papapana to Tok Pisin but not to any local language: 66% of the community speak Tok Pisin as a L1, 17% speak Papapana as a L1 and the local language with the highest representation of L1 speakers is the Papuan language Rotokas with 4% (see Table 3).

There has been a significant increase in contact between Papapana speakers and people from different linguistic backgrounds over the past century, and subsequently there has been more need for and exposure to a lingua franca like Tok Pisin. Nevertheless, this is not a reason in itself to abandon Papapana. The Papapana speech community have historically been multilingual, so why not add Tok Pisin to their linguistic repertoire while maintaining Papapana? What has changed to bring about the decreased efficacy of Papapana and result in its abandonment?

The first factor is the community's coastal location. Coast-dwellers are more likely than mountain-dwellers to have contact with outsiders who arrive by boat. Since most of the plantations and mission stations in Bougainville were established

along the coast, Papapana people were not only more likely to have initial contact with outsiders, but this contact was more sustained. During the Bougainville Crisis, coastal people would also have been more vulnerable to the invading PNG Defence Force and therefore more likely to be displaced from their homes, either being placed in Care Centres or fleeing to the mountains.

Secondly, the Papapana speech community has always been a minority immigrant group. While migration, immigration and increased mobility to population centres do not necessarily lead to shift, they can have a significant impact if coupled with a small population, causing a noticeable and more rapid decrease in the overall number and proportion of speakers. In the Papapana community, the demographic changes have reduced the efficacy of Papapana as there are fewer interlocutors and thus fewer opportunities to use the language, and immigrants have less opportunity and motivation to learn Papapana.

In addition, Papapana is not viewed as a powerful language because of its small speaker numbers. The demographic changes in the Papapana community have also likely affected their sense of identity as community ties have been weakened. Community members expressed the importance of Papapana to their culture and a wish for its continued use. Just as Kulick (1992) found for Taiap speakers in Papua New Guinea, Tok Pisin was viewed by Papapana speakers as belonging to everyone, and being a Melanesian language, and it was valued for its role as a lingua franca; however some viewed it as "bad English". Nevertheless, speakers are still shifting to Tok Pisin, which suggests that either the positive views of Papapana are not entirely accurate, or that on a subconscious level, there has been a shift from ethnic/community identity to regional/national identity; a change seen elsewhere in PNG after national independence (Wurm 2007: 444).

The third factor is the lack of representation in various domains. This does not necessarily lead to shift, since a state of diglossia could exist; however, coupled with a decreasing speaker base, the lack of, or weak, representation of Papapana in church, education and the media has contributed significantly to shift since it adds to the occasions in which Papapana is not used and affects its perceived importance, which negatively impacts speakers' motivations to use Papapana. Speaker demographics also mean Papapana is less likely to be represented in these institutions: there is less chance of there being Papapana-speaking teachers and students, and there is less chance of a bible translation organisation working with the community because there are not enough speakers to make a bible translation logistically possible. Since children are not usually exposed to the more complex grammatical structures of a language until school, then a lack of continued literacy education in Papapana means opportunities to acquire writing and formal styles of expression in Papapana never exist and Papapana also loses prestige and usefulness in domains that require such skills.

The fourth factor is that the socioeconomic status and motivation of the Papapana speech community is seemingly low following the Bougainville Crisis: some outsiders expressed the impression that the Papapana villages are slower to recover from the Crisis than other Bougainville villages. It is therefore possible that Papapana speakers associate their current lifestyle with their language, thus decreasing its prestige. Papapana is also not seen as a vital language and while this inspired two speakers to make a conscious effort to transmit Papapana to their two children, for others it may be a self-fulfilling prophecy: community members may feel that it is not beneficial to speak or transmit Papapana as it has no future.

The interaction of all these factors has led to changes in Papapana speakers' attitudes towards their language: Papapana has lost prestige and usefulness and is therefore being abandoned. These factors are all interwoven in a complex chain of causation, making it impossible to single out one factor as the primary cause of language shift.

4.   Assessing ethnolinguistic vitality status

There are a multitude of sociolinguistic assessment frameworks that seek to label a language in terms of its vitality, including Fishman's (1991) Graded Intergenerational Disruption Scale (GIDS) (§4.1), UNESCO's Language Vitality and Endangerment framework (Brenzinger et al. 2003) (§4.2), Landweer's (2012) Indicators of Ethnolinguistic Vitality (IEV) (§4.3) and the Catalogue of Endangered Languages Project's Language Endangerment Index (LEI) (Lee & Way 2016) (§4.4). This section assesses Papapana using these frameworks and evaluates the frameworks on their applicability and whether they make the distinction between "directly observable symptoms (indications) of language endangerment and their often not so clearly discernible causes" (Himmelmann 2010: 46): "symptoms" indicate the degree of language vitality/endangerment, while "causes" are factors which have led, or may lead, to weak ethnolinguistic vitality. This evaluation is elaborated in §4.5.

4.1   Fishman's (1991) GIDS

From the 1990s, attention focused on intergenerational language transmission as a key factor in linguistic vitality and a number of categorisation scales were created including those of Krauss (1997: 25–26, 2007: 1) and Wurm (1998: 192). The best-known and most influential intergenerational transmission scale is Fishman's (1991) Graded Intergenerational Disruption Scale (GIDS) (Table 6),

Table 6. Fishman's (1991) Graded Intergenerational Disruption Scale (GIDS) (adapted from Fishman 1991; Lewis & Simons 2010)

Level	Description
1	The language is used in education, work, mass media, government at the nationwide level
2	The language is used for local and regional mass media and governmental services
3	The language is used for local and regional work by both insiders and outsiders
4	Literacy in the language is transmitted through education
5	The language is used orally by all generations and is effectively used in written form throughout the community
6	The language is used orally by all generations and is being learned by children as their first language
7	The child-bearing generation knows the language well enough to use it with their elders but is not transmitting it to their children
8	The only remaining speakers of the language are members of the grandparent generation

which focuses on intergenerational transmission as a key factor in the maintenance of a language (levels 6–8), but also considers domains (levels 1–3) and literacy (levels 4–5) since societal and institutional choices are crucial in influencing parental decisions about language choice. In the first six levels of GIDS, the language is being maintained, while in seven and eight, intergenerational transmission has ceased and language shift has begun.

On the GIDS, Papapana is at level 4 since some literacy, albeit limited, is transmitted through education (see §2.3.2), but it not clear what degree of literacy this criterion refers to. Papapana can also be classified as between level 6 and 7 since there are speakers in each generation who speak the language but with the exception of one couple, the child-bearing generation are not transmitting Papapana to their children (§2.3.1). Since the GIDS considers different factors at different levels within one categorisation scale, more than one level applies to Papapana. To choose a particular level, as the *Ethnologue* (Lewis et al. 2014) does for Papapana, would be to ignore other pertinent factors and thus be misleading.

4.2 UNESCO (2003) Language vitality and endangerment framework

At the International Expert Meeting on UNESCO Programme Safeguarding of Endangered Languages, Brenzinger et al. (2003) proposed nine factors that can together be used to characterise a language's overall sociolinguistic situation. Brenzinger et al. (2003) identified six factors that can be used to evaluate a language's vitality and state of endangerment, two factors to assess language attitudes

and one factor to evaluate the urgency for documentation. It should be noted that in application these nine factors are sometimes all falsely assumed to be indicators, as in Lewis (2006). Like Fishman's (1991) GIDS, the UNESCO framework considers intergenerational transmission, domains, and literacy and education materials; however, the UNESCO framework treats each of these factors on distinct six-levelled scales. The UNESCO framework also considers response to new domains and both the absolute and relative population of speakers as factors in assessing linguistic vitality.

For the first of the six factors, intergenerational language transmission, Papapana scores grade 3 "definitively endangered" as it "is used mostly by the parental generation and up" (Brenzinger et al. 2003: 8) (see §2.3.1). For the second factor, absolute number of speakers, there is no scale but we can note that Papapana has 106 fluent speakers (§2.1). Landweer (2012) and ELCat's LEI (Lee & Way 2016) also consider the number of speakers as an indicator of linguistic vitality: a language with large speaker numbers is assumed to be more vital than a language with small speaker numbers. However, establishing accurate figures is problematic. Firstly, sources do not always distinguish between the ethnic group's population and the number of speakers of its traditional language (Kibrick 1991). Secondly, the definition of *speaker* is debatable. Thirdly, even if one decides on a level of proficiency that is required to qualify as a speaker, how does one go about testing every individual's proficiency in a large population? Furthermore, even if an accurate number of speakers were established, there is no consensus in the literature as to what constitutes a "safe" figure: figures range from 100 to 10,000 to 20,000 speakers (see Crystal 2000: 12; Dixon 1991: 231). For Melanesia, such figures are unhelpful since Melanesian languages have an average of 2,382 speakers (Landweer 2012: 153). More importantly, while it is true that a small speech community is more vulnerable to decimation or merging with a neighbouring group (Brenzinger et al. 2003: 8), even a language with large speaker numbers could be in danger "if the external pressures on it were great, while a very small language could be perfectly safe as long as the community was functional and the environment stable" (Nettle & Romaine 2000: 41). For example, Breton (France) has 280,000 speakers but is considered to be facing extinction, while Gumawana (PNG) has 367 speakers but displays great vitality (Barrena et al. 2007: 135–136). Absolute speaker numbers does not therefore indicate the extent to which a language is endangered.

The third UNESCO factor is the proportion of speakers within the total population. Fluent Papapana speakers only account for between 17% and 21% of the total population of the Papapana villages (§2.1). If this is considered a "minority" then Papapana scores grade 2, "severely endangered", though if this is considered

"very few", Papapana scores grade 1, "critically endangered" (Brenzinger et al. 2003:9). The UNESCO level descriptors are thus a little vague for this factor.

The fourth factor concerns trends in existing language domains. As §2.3 reports, Tok Pisin is the language of all domains, but Papapana could be used among Papapana speakers in these domains, and while Papapana exists most strongly in the home domain, the dominant language is increasingly Tok Pisin. Papapana thus seems to fall between grade 2 and 3 for this factor: at grade 2, "the language is used in limited social domains and for several functions", while at grade 3, "the language is in home domains and for many functions, but the dominant language begins to penetrate even home domains" (Brenzinger et al. 2003:10).UNESCO's fifth factor is response to new domains (such as work/educational environments) and media (including broadcast media and the Internet) (Brenzinger et al. 2003:11). Papapana has not been adopted in any new domains (§2.3) so it scores zero "inactive".

The sixth factor concerns materials for language education and literacy, but for each level of the scale, a number of variables are considered (Brenzinger et al. 2003:12) and consequently Papapana falls between grade 1, where "a practical orthography is known to the community" (although only a few speakers are literate) and "some material is being written" (albeit not permanent materials), and grade 3, where "children may be exposed to the written form at school. Literacy is not promoted through print media" (although school literacy is limited). This factor should arguably not be considered an indicator of linguistic vitality. A lack of materials may reflect institutional or community attitudes, and/or it might mean language use in the school domain is undermined; both of these things can cause language shift. Nevertheless, educational and literacy materials are not a directly observable symptom of linguistic vitality; if they were, that presupposes that every culture has a school domain and wishes their language to be written, which is not always the case. Using the accessibility of materials as evidence for language use in the school domain is also problematic: the existence of materials does not mean the language *is* used in the school domain, and conversely a language could still be used even if there were no materials.

The UNESCO framework identifies Papapana as endangered to varying degrees. By separating the factors it is possible to see that what makes Papapana most vulnerable is the low proportion of speakers and the fact that Papapana is not used in new domains.

4.3   Landweer (2012) Indicators of Ethnolinguistic Vitality (IEV)

Landweer (2012: 153) argues that previous assessment tools have been "developed in sociolinguistic contexts that do not coincide with Melanesia". Indeed, Landweer (2012: 161) found that 97.6% of the language case studies cited in benchmark theoretical and typological constructs were outside of Melanesia. In response to this and taking into account the ways in which the Melanesian sociolinguistic context is different, Landweer (2012: 164–170) proposes eight Indicators of Ethnolinguistic Vitality (IEV). Each of these indicators has a score of 0–4 and these points are accumulated to give an overall vitality score.

Indicator 1 is potential for contact and Papapana scores 1 since, as described in §2.2, the speech community has "fairly easy access to and from the nearest urban (or population) centre" (Landweer 2012: 164), which can include government stations, regional schools and mission compounds. Landweer (2012: 164) argues that access should take into account the availability and cost of transport as well as the distance, since the perception of remoteness is culturally defined and in PNG people regularly travel several hours to and from their villages. Language contact is of course a factor in language endangerment, precisely because it is a prerequisite for language shift. Identifying the type of language contact might help in predicting the likelihood of endangerment or the type of shift to occur. Nevertheless, language contact does not necessarily lead to language shift and is therefore not an indicator of endangerment.

Indicator 2 concerns domains of usage, including the home, "traditional" cultural events and "Western" social events (Landweer 2012: 165). Based on the information in §2.3, Papapana scores 0 as it is "mixed with a lingua franca or other languages in every domain across society include [sic] the home environment" (Landweer 2012: 165).

Indicator 3 considers frequency and type of code-switching, which Landweer (2012: 166) defines as occurring "when a speaker embeds elements from one language in an utterance that is primarily composed of another language". Landweer (2012: 166) differentiates "inter-sentential code-switching", which occurs at major communication boundaries, and "intra-sentential code switching", which occurs within a single thought group and typically without redefinition of the communicative situation, and can thus be referred to as "unbounded". Landweer claims that the more frequent intra-sentential code-switches are, the more endangered the language is. Impressionistically, Papapana scores 0 for this indicator as there is "frequent individual unbounded code switching" (Landweer 2012: 166); however, the relationship between code-switching and language endangerment needs extensive investigation before it can be factored into an ethnolinguistic vitality assessment framework.

Indicator 4, 'population and group dynamics', relates to the need for a core of fluent speakers which can be supported or undermined by the language use of immigrants. Papapana scores between 1 and 2 for this indicator since "immigrants require two-way communication entirely via a lingua franca" (score 1), but some immigrants are passively bilingual (score 2) (Landweer 2012: 167).

Indicator 5 is based on Milroy's (1987) concept of social networks: Landweer (2012: 167) argues that "dense and multiplex networks can serve to insulate speakers, isolating and protecting them from language contact pressures towards change". Based on the information presented in §2.2.6 and §2.3, Papapana best fits a score of 1: "divided network systems, internally dense, however, there is the necessity to communicate with outsiders who do not know the local language for all goods and services" (Landweer 2012: 167). Since this factor relates to language contact, of which language shift is just one possible outcome, social networks is also arguably not an indicator of vitality.

Indicator 6, 'social outlook', questions whether there is internal and/or external recognition of the language community as separate and unique within the broader society, and whether there is material or non-material evidence (cultural markers) of such a distinction. Based on the discussion in §3.2, Papapana scores 1: "weak internal identity, neutral status conferred by outsiders, with some cultural markers present" (Landweer 2012: 168). The UNESCO framework (Brenzinger et al. 2003) also includes community members' attitudes towards their language as a factor, though crucially it does not include this factor as an indicator. Instead UNESCO recognises that community attitudes "interact with governmental policy and societal pressures *to result in* increased or decreased language use in different domains" (Brenzinger et al. 2003: 15, emphasis added). Indeed, it is a change in attitudes, brought about by environmental changes, that is the ultimate cause of language shift.

Indicator 7 is language prestige. Based on the discussion in §3.2, Papapana scores 1 as it is "a locally recognized variety with neutral status" (Landweer 2012: 169). It is true that many safe languages enjoy official status within nations and consequently tend to be held in higher prestige (Grenoble and Whaley 2006: 18), while endangered languages tend not to; however, there are many languages that are safe even though they do not have official status. This is particularly the case in multilingual nations with great linguistic diversity, such as PNG. Conversely, equal legal status does not guarantee language maintenance and long-term vitality (Brenzinger et al. 2003: 13). Moreover, language status and policies reflect institutional attitudes towards a language which are among the *causes* of language shift and endangerment. Indeed, although institutional attitudes and policies is one of UNESCO's nine factors, it is not one of their six indicators of linguistic vitality

(Brenzinger et al. 2003): "the linguistic ideology of a state may *inspire* linguistic minorities to mobilize their populations toward the maintenance of their languages, or may *force* them to abandon them" (Brenzinger et al. 2003: 12, emphasis added).

Indicator 8 concerns access to a stable and acceptable economic base. Based on the discussion in §2.3.3 Papapana scores 1: "marginal subsistence economy requiring augmentation of the traditional means of subsistence with cash-based economic schemes requiring use of a language other than the target language" (Landweer 2012: 169). Landweer (2012: 169) is right that "one of the most common motivations for individuals in a community to shift from one language to another is for perceived economic benefit"; however, Landweer (2012) calls this factor an "indicator" yet uses the term "motivation" to describe it, which is exactly what this factor is. Whether or not there is access to a stable and acceptable economic base tells us nothing about linguistic vitality or language use; it tells us about a possible cause for speakers' attitudes which in turn motivate language choice and use.

Overall, Landweer's (2012) IEV is well defined. Cumulatively, Papapana scores between 6 and 7 and can be labelled "endangered": a score of 18–24 points indicates "probable continued language viability", 15–17 points indicates "possible viability", 12–14 is considered "on the cusp of shifting" and 0–12 is "endangered" (Landweer 2012: 163). These margins of six, three and twelve points do however seem rather inconsistent and consequently the scale seesaws between broad- and fine-grained. Most troubling is that many of Landweer's "indicators" are actually potential causes of language shift rather than measures of ethnolinguistic vitality.

4.4 Catalogue of Endangered Languages Project's (ELCat) Language Endangerment Index (LEI) (Lee & Way 2016)

The most recent assessment framework is The Catalogue of Endangered Languages Project's (ELCat) Language Endangerment Index (LEI) (Lee & Way 2016). The LEI considers four categories: intergenerational transmission, absolute speaker numbers, speaker number trends and domains of use. For each category, a language is assigned a score of 0–5 based on how well it meets the criteria shown in Table 7. Intergenerational transmission is worth twice each of the other factors.

Papapana scores 2–3 for intergenerational transmission, which when doubled gives a score of 4–6: the criteria for the score of 2 fits Papapana nicely as it allows for the fact that children are "generally" not speakers, however, the conflict here is that "some adults" (from score 3) is more accurate than "most adults" (score 2). A score of 3 is assigned for absolute speaker numbers, 4 for speaker trends and 4 for domains; however, score 0 for domains also applies because Papapana is used

Table 7. Catalogue of Endangered Languages Project's Language Endangerment Index (LEI) (adapted from Lee & Way 2016)

Level of endangerment	5 Critically endangered	4 Severely endangered	3 Endangered	2 Threatened	1 Vulnerable	0 Safe
Intergenerational Transmission	Only a few elderly speakers.	Many grandparents speak the language; younger people generally do not.	Some adults are speakers; children are not.	Most adults are speakers; children generally are not.	Most adults and some children are speakers.	All members of the community speak the language.
Absolute speaker numbers	1–9 speakers	10–99 speakers	100–999 speakers	1000–9999 speakers	10,000–99,999 speakers	>100,00 speakers
Speaker number trends	A small percentage of the community speaks the language; numbers are decreasing very rapidly.	Less than half of the community speaks the language; numbers are decreasing at an accelerated pace.	About half of community members speak the language; numbers are decreasing steadily.	A majority of community members speak the language; numbers are gradually decreasing.	Most members of the community speak the language; numbers may be decreasing, but very slowly.	Almost all community members speak the language; numbers are stable or increasing.
Domains of use	Used only in a few very specific domains, such as in ceremonies, songs, prayer, proverbs, or certain limited domestic activities.	Used mainly in the home and/or with family; may not be the primary language even in these domains for many community members.	Used mainly in the home and/or with family; remains the primary language of these domains for many.	Used in some non-official domains along with other languages; remains the primary language used in the home for many.	Used in most domains except for official ones such as government, mass media, education etc.	Used in most domains, including official ones.

in school, though overall score 4 fits best as Papapana is certainly not used "in most domains".

If no information is available for a particular category, it is not scored for that category and the number of points available from that category is deducted from the total points available. Since there is information available for all categories for Papapana, the total points available are 25. A percentage is then calculated from the total score and the total points available: 0% indicates "safe", 1–20% "vulnerable", 21–40% "threatened", 41-60% "endangered", 61–80% "severely endangered", and 81–100% "critically endangered". Papapana scores between 64% or 72% and is thus categorised "severely endangered". A level of certainty accompanies each endangerment score, showing the degree of confidence in the score and is calculated based on the percentage of factors that are known and entered. For Papapana, all factors are known so the certainty level is 25/25, i.e. 100%. Languages that score 0% but have a certainty level of less than 100% are identified as "at risk". The catalogue also identifies some languages as "dormant" if a source reports that there are no known L1 speakers, or as "awakening" if the language is being revitalised.

4.5 Evaluation of assessment frameworks

Although the assessment frameworks all show that Papapana is endangered, there is no consensus on the degree to which it is endangered. Some of the problems with the assessment frameworks are vague definitional criteria and the combination of factors on one level. The UNESCO framework (Brenzinger et al. 2003), Landweer's (2012) IEV and ELCat's LEI (Lee & Way 2016) are better designed than the GIDS (Fishman 1991) because they separate individual components of the assessment, allowing a better understanding of where weaknesses and strengths lay. Their downfall, however, could be that they require in-situ studies, which might not always be possible. Indeed, Lewis (2006: 28–29) applied the UNESCO framework to the world's languages but concluded that more data and better reporting were needed to make such an assessment. Assessing Papapana emphasised the need for in-situ observations as the vernacular education policy did not reflect actual educational practice, and without participant observation, the two Papapana-speaking children may have gone unnoticed. An assessment framework should therefore allow for the intricacies that an in-situ assessment reveals, but not depend on it. Indeed, individuals do not shift languages to the same extent at the same time, and languages do not completely disappear from domains in a uniform order; therefore there is a need to allow for exceptions. A good assessment framework should also be relevant to a variety of language shift situations and should not assume, like the frameworks discussed here, that all shift is top-down.

A major problem is that many of the assessment frameworks did not make a distinction between "symptoms" and "causes" of language endangerment (Himmelmann 2010: 46). Admittedly, it is very difficult to untangle indicators and causes of language endangerment. This study has shown that proportion of speakers within a community can indicate linguistic vitality status; however, demographic changes in absolute and relative number of speakers can also be a cause. Similarly, while the number and quality of domains is an indicator of vitality, changes in domain usage can cause attitudes to change and language shift to occur. Sometimes an indicator might provide crucial clues to the causes of shift. For instance, with two exceptions, the youngest fluent Papapana speakers were born in the early to mid-1980s. It is unlikely to be a coincidence that the Bougainville Crisis began in the late 1980s, causing population displacement and contributing to language shift. Despite the difficulty, it is nevertheless possible to untangle symptoms and causes. The factors which emerge as symptoms/indicators of language endangerment are proportion of speakers, and domains of language use, with intergenerational transmission relating to the home domain.

The UNESCO framework (Brenzinger et al. 2003) and ELCat's LEI (Lee & Way 2016) consider the proportion of speakers as an indicator of linguistic vitality. This indicator is perhaps more problematic than absolute numbers with regards to establishing accurate figures since one needs to not only gather figures on speaker numbers, but also population figures for the speakers' communities. One still also needs to define *speaker* and accurately assess speaker's proficiency (see §4.2). Nonetheless, the proportion of speakers is much more meaningful than absolute numbers. As Bauer (2008: 63) explains, for there to be a reasonable chance that a language will be spoken, there has to be a reasonable chance that those spoken to will also speak the language. If speakers are too diluted in their community by non-speakers, then they will not have many opportunities to speak the language and the conversation is more likely to shift to a lingua franca.

The domains in which a language is used is considered an indicator of linguistic vitality by all the frameworks. The fewer domains a language is spoken in, the less opportunity speakers have to speak their language and its use is less reinforced and maintained. Himmelmann (2010: 46) argues that the number and quality of domains is "one essential symptom for the vitality of a language". The number of domains is important because different domains involve different registers, but the quality of domains is also significant, that is, "the importance of a given domain within the overall language ecology in a given speech community, based on the breadth and variety of linguistic behaviour found in that domain" (Himmelmann 2010: 46). For example, the use of a language once a week in church could be seen as less relevant to linguistic vitality than the use of the language every day in school. I would also add that it is important to consider the

dominance of a language within a domain, and ELCat's LEI (Lee & Way 2016) goes some way to addressing this in its inclusion of which language is "primary" in a domain. Investigating language use in particular domains involves in-situ observations if one is to gain reliable information. For example, Papapana community members might say that Papapana is used in church but in-situ observations showed that this really only applied to the occasional song, while PNG language policy might state that the vernacular language is to be used as the medium of instruction in elementary schools but in-situ observations revealed that Papapana was far from the language of education.

School education is one type of domain that might exist in a community and is considered by several of the frameworks. If a language is not spoken in the school domain it does not necessarily indicate that the language is endangered; the school domain should be considered along with other domains when assessing linguistic vitality.

All the frameworks, except Landweer's (2012) IEV, considered intergenerational language transmission as a crucial indicator of linguistic vitality. As with any domain, identifying language use in the home may involve observations, particularly of child-adult interactions. Home domain usage can also be identified by measuring intergenerational language transmission. The parents may choose, consciously or subconsciously, not to transmit their language to their children, or it could be that the parents are transmitting but the children choose to respond in another language. These choices would be the result of a change in attitudes brought about by an environmental change. The directly observable symptom is an interruption in language transmission. The home domain is arguably the most important domain, since the interaction between caregivers and children determines the existence/absence of intergenerational language transmission, and consequently the fate of the language because as speakers die out, the speaker base is not replenished.

Proportion of speakers and domains of language use must be considered together to provide an accurate assessment of ethnolinguistic vitality: a language could exist in a particular domain but if there is a low proportion of speakers, it might not be used much in that domain.

5. **Papapana speakers in transition: The hows and whys of language endangerment**

Papapana is at great risk of disappearing within the next century due to voluntary, top-down shift to Tok Pisin. Economic and cultural changes have increased mobility to population centres which has increased contact between Papapana

speakers and other speech communities. In turn, intermarriage has increased, which has further increased contact, both inside and outside of the Papapana villages. The Papapana community has undergone further population movement due to permanent migration, and displacement from colonial expansion and the Bougainville civil war, and has been particularly vulnerable to contact due to the community's coastal location. Increased contact among people with such diverse linguistic backgrounds has heightened the exposure and need for the lingua franca Tok Pisin. The weak representation in institutions such as school, church and the media has resulted in the use and prestige of Papapana being undermined in these domains. These factors combined with an already small speaker base have left the population of Papapana speakers smaller than ever before. The small proportion of speakers and the lack of support in particular domains, means there are fewer opportunities to speak Papapana, and speakers' attitudes towards Papapana are seriously damaged. A change in attitudes may also be attributed to the perception of Papapana as a weak, endangered language. It is also possible that subconsciously there has been a shift from ethnic identity to regional identity which has promoted the use of Tok Pisin even further.

Papapana is now spoken by less than 20% of the total population of the villages where it was traditionally spoken, intergenerational transmission has almost ceased with only two children speaking Papapana as a L1 and Tok Pisin is the dominant language of all domains, though Papapana may be used among Papapana speakers in these domains and is used in elementary school to a limited degree. It is safe to say that Papapana is "endangered" but unfortunately the assessment frameworks described here did not further delineate Papapana's vitality status. Instead the assessment of Papapana revealed some of the problems with these frameworks which should be addressed in the development of future models: a lack of clear definitional criteria, the combination of different factors onto one level, the need for in-situ studies, and the false assumption that individuals shift languages to the same extent at the same time. The biggest issue was the lack of distinction between symptoms and causes of language endangerment. A dead language is one which is not being used anymore; therefore, if we want to assess how close to that point of death a language is, i.e. the extent of endangerment, we need to assess how much the language is being used. Usage depends on opportunities to use the language in particular settings and with particular interlocutors. The number and quality of domains, and the proportion of speakers within a community are thus crucial indicators of linguistic vitality: the more domains and the higher the proportion of speakers, the greater the chance of the language being used. Intergenerational transmission is also an indicator as it reflects the extent to which the language is used in the home and school domain, and indicates

whether the speaker base is being replenished for the future. Other factors such as the potential for contact, literacy materials and language attitudes are potential causes of language endangerment and belong to a predictive model, that is, one which identifies the vulnerability of a language to endangerment and predicts the fate of a language. It is important to make a clear distinction between causes and symptoms as such assessments might have different applications: a predictive model may be used to prevent or reverse the demise of a language, while a diagnostic model may be used to identify and prioritise the languages that most need documenting or revitalising.

References

Baker, P. & Mühlhäusler, P. 1996. The origins and diffusion of Pidgin English in the Pacific. In *Atlas of Languages of Intercultural Communication in the Pacific, Asia, and the Americas*, Vol II(1), S.A. Wurm, P. Mühlhäusler & D.T. Tryon (eds), 551–583. Berlin: Mouton de Gruyter.

Barrena, A., Amorrortu, E., Ortega, A., Uranga, B., Izagirre, E. & Idiazabal, I. 2007. Does the number of speakers of a language determine its fate? *International Journal of the Sociology of Language* 186: 125–139.

Bauer, W. 2008. Is the health of Te Reo Māori improving? *Te Reo* 51: 33–73.

Brenzinger, M., Yamamoto, A., Aikawa, N., Koundiouba, D., Minasyan, A., Dwyer, A., Grinevald, C., Krauss, M., Miyaoka, O., Sakiyama, P., Smeets, R. & Zepeda, O. 2003. Language vitality and endangerment. Document submitted to the *International Expert Meeting on UNESCO Programme Safeguarding of Endangered Languages*, Paris, 10–12 March. <http://www.unesco.org/culture/ich/doc/src/00120-EN.pdf>

Campbell, L. & Muntzel, M.C. 1989. The structural consequences of language death. In *Investigating Obsolescence: Studies in Language Contraction and Death*, N.C. Dorian (ed.), 181–196. Cambridge: CUP. doi:10.1017/CBO9780511620997.016

Crystal, D. 2000. *Language Death*. Cambridge: CUP. doi:10.1017/CBO9781139106856

Dixon, R.M.W. 1991. The endangered languages of Australia, Indonesia and Oceania. In *Endangered Languages*, R.H. Robins & E.M. Uhlenbeck (eds), 229–255. Oxford: Berg.

Evans, B. 2009. Beyond pronouns: Further evidence for South Bougainville. In *Discovering History through Language: Papers in Honour of Malcolm Ross*, B. Evans (ed.), 73–101. Canberra: Pacific Linguistics.

Fishman, J.A. 1991. *Reversing Language Shift*. Clevendon: Multilingual Matters.

Florey, M. 2005. Language shift and endangerment. In *The Austronesian Languages of Asia and Madagascar*, K.A. Adelaar & N. P. Himmelmann (eds), 43–64. London: Routledge.

Foley, W.A. 2004. Language endangerment, language documentation and capacity building: challenges from New Guinea. In *Language Documentation and Description*, P.K. Austin (ed.), 28–38. London: Hans Rausing Endangered Languages Project, Department of Linguistics.

Grenoble, L.A. & Whaley, L.J. 1998. Toward a typology of language endangerment. In *Endangered Languages: Current Issues and Future Prospects*, L.A. Grenoble & L.J. Whaley (eds), 22–54. Cambridge: CUP.

Grenoble, L.A. & Whaley, L.J. 2006. *Savings Languages: An Introduction to Language Revitalization*. Cambridge: CUP.

Himmelmann, N.P. 2010. Language endangerment scenarios: A case study from Northern Central Sulawesi. In *Endangered Languages of Austronesia*, M. Florey (ed.), 45–72. Oxford: OUP.

Kibrik, A.E. 1991. The problem of endangered languages in the USSR. In R.H. Robins & E.M. Uhlenbeck (eds), *Endangered Languages*, 257–273. Oxford: Berg.

Krauss, M. 1992. The world's languages in crisis. *Language* 68: 4–10. doi:10.1353/lan.1992.0075

Krauss, M. 1997. The indigenous languages of the North: A report on their present state. Northern minority languages: problems of survival. *Senri Ethnological Studies* 44: 1–34.

Krauss, M. 2007. Classification and terminology for degrees of language endangerment. In *Language Diversity Endangered*, M. Brenzinger (ed.), 1–8. Berlin: Walter de Gruyter.

Kulick, D. 1992. *Language Shift and Cultural Reproduction: Socialization, Self, and Syncretism in a Papua New Guinean Village*. Cambridge: CUP.

Landweer, M.L. 2012. Methods of language endangerment research: A perspective from Melanesia. *International Journal of the Sociology of Language* 214: 153–178.

Laracy, H. 1976. *Marists and Melanesians: A History of Catholic Missions in the Solomon Islands*. Honolulu HI: The University Press of Hawai'i.

Laracy, H. 2005a. 1914: Changing the guard at Kieta. In *Bougainville Before the Conflict*, A.J. Reagan & H.M. Griffin (eds), 136–140. Canberra: Pandanus Books.

Laracy, H. 2005b. '*Imperium in Imperio*'? The catholic church in Bougainville. In *Bougainville Before the Conflict*, A.J. Reagan & H.M. Griffin (eds), 125–135. Canberra: Pandanus Books,

Lee, N.H. & Way, J.V. 2016. Assessing levels of endangerment in the Catalogue of Endangered Languages (ELCat) using the Language Endangerment Index (LEI). *Language in Society* 45(2): 271–292.

Lewis, M.P. 2006. Towards a categorization of endangerment of the world's languages. *SIL Electronic Working Papers 2006*-002: 31.

Lewis, M.P. & Simons, G.F. 2010. Assessing endangerment: Expanding Fishman's GIDS. *Revue Roumaine de Linguistique* 2: 103–120.

Lewis, M.P., Simons, G.F. & Fennig, C.D. 2014. *Ethnologue: Languages of the World*. 17th edn. Dallas TX: SIL International.

Lynch, J. 1998. *Pacific Languages: An Introduction*. Honolulu HI: University of Hawai'i Press.

Lynch, J., Ross, M. & Crowley, T. (eds). 2002. *The Oceanic Languages*. Richmond: Curzon Press.

Matras, Y. 2005. Language contact, language endangerment, and the role of the 'salvation linguist'. In *Language Documentation and Description*, P.K. Austin (ed.), 225–251. London: Hans Rausing Endangered Languages Project, Department of Linguistics, SOAS.

McHardy, E. 1935. *Blazing the Trail in the Solomons*. Sydney: Visitor Printing Company.

Mühlhäusler, P. 1976. Samoan Plantation Pidgin and the origins of New Guinea Pidgin: An introduction. *Journal of Pacific History* 11(2): 122–125. doi:10.1080/00223347608572295

Milroy, L. 1987. *Language and Social Networks*. Oxford: Basil Blackwell.

NDOE. 2003. *Elementary Language Syllabus*. Waigani: National Department of Education (NDOE).

Nettle, D. & Romaine, S. 2000. *Vanishing Voices: The Extinction of the World's Languages*. Oxford: OUP.

Oliver, D. 1991. *Black Islanders: A Personal Perspective of Bougainville 1937–1991*. Melbourne: Hyland House.

Regan, A.J. & Griffin, H.M. (eds). 2005. *Bougainville before the Conflict*. Canberra: Pandanus Books.

Ross, M. 1996. Mission and church languages in Papua New Guinea. In *Atlas of Languages of Intercultural Communication in the Pacific, Asia, and the Americas*, Vol II, S.A. Wurm, P. Mühlhäusler & D.T. Tryon (eds), 595–617. Berlin: Mouton de Gruyter.

Tryon, D. 2005. The languages of Bougainville. In *Bougainville Before the Conflict*, A.J. Reagan & H.M. Griffin (eds), 31–46. Canberra: Pandanus Books.

Vernon, D. 2005. The Panguna mine. In *Bougainville Before the Conflict*, A.J. Reagan & H.M. Griffin (eds), 258–273. Canberra: Pandanus Books.

Winford, D. 2003. *An Introduction to Contact Linguistics*. Oxford: Blackwell.

Wurm, S.A. 1979. The language situation in the New Guinea area. In *New Guinea and Neighboring Areas: A Sociolinguistic Laboratory*, S.A. Wurm (ed.), 3–10. Paris: Mouton.

Wurm, S.A. 1998. Methods of language maintenance and revival, with selected cases of language endangerment in the world. In *Studies in Endangered Languages: Papers from the International Symposium on Endangered Languages*, *Tokyo* Nov 18–20 1995, K. Matsumura (ed.). Tokyo: Hituzi Syobo.

Wurm, S.A. 2003. The language situation and language endangerment in the Greater Pacific area. In *Language Death and Language Maintenance: Theoretical, Practical and Descriptive Approaches* [Current Issues in Linguistic Theory 240], M. Janse & S. Tol (eds), 15–47. Amsterdam: John Benjamins. doi:10.1075/cilt.240.04wur

Wurm, S.A. 2007. Australasia and the Pacific. In *Encyclopedia of the World's Endangered Languages*, C. Moseley (ed.), 425–466. Oxon: Routledge.

Language transmission
Shift, loss and survival

The art of losing

Beyond *java*, *patois* and postvernacular vitality – Repositioning the periphery in global Asian ecologies

Lisa Lim
University of Hong Kong

This paper discusses issues in endangerment and postvernacularity in the context of Asia, a region with complex dynamics in multilingual ecologies that also includes the presence – dominance – of English, a language that entered the ecologies through colonisation. I use as illustration two minority communities with endangered vernaculars – the Malays of Sri Lanka, brought from various parts of the Malay archipelago by the Dutch and British colonial powers, and their vernacular Sri Lanka Malay, traditionally known as *java*, a mixed language of trilingual base (Malay, Sinhala, Tamil); and the Peranakans, descendants of southern Chinese merchants who settled in Malaya and intermarried with local women, and their vernacular Baba Malay, a restructured variety of Malay with southern Sinitic influences, usually referred to as *patois*. I query if linguistic and cultural loss is inevitable, or if such situations of shift – to a language of wider communication or an emergent variety – are in fact instances of empowerment and evolution in response to change, where a repositioning of the periphery in the new global economy brings greater accessibility to and participation in the Centre, and better adaptation for surviving and thriving.

Keywords: Asia, ecology, empowerment, patois, Peranakans, postvernacular, shift, Sri Lanka Malay

1.   Introduction

"The art of losing", says poet Elizabeth Bishop (1983), "isn't hard to master / so many things seem filled with the intent / to be lost that their loss is no disaster". The potential loss of 90% of the world's languages within 50 to 100 years is, in the discourse of endangerment, by now widely recognised, and has been met

in recent decades with urgent calls and initiatives by scholars and supranational bodies such as UNESCO to counter such an inevitability. But a traditional Peranakan saying suggests: *Dah sa chupak tak boley sa gantang* 'a quart will never make a gallon', that is: you cannot change destiny. In this paper[1] I contemplate these sentiments: Is the loss of the ancestral language and culture of a minority community indeed inevitable? What factors play a part in the art of losing such that the destiny of such communities evolves to become empowerment rather than endangerment?

I discuss the issues involved in situations of shift and endangerment in the context of Asia, a region that affords complex dynamics in multilingual ecologies that also includes the presence – dominance – of English, a language that entered the ecologies during colonisation. To this end I use in a close study two communities that are interesting to examine for points of comparison and contrast: the Malays of Sri Lanka, and the Peranakans of Singapore. Both were formed in the time of, and as a consequence of, the development of exploitation colonies in Asia – globalisation of another age. Both are considered creole communities, and have seen language shift and endangerment of their ancestral languages – respectively Sri Lanka Malay, known to the community as *java*, and Baba Malay, referred to as *patois* – and revitalisation. Additionally significant is that these are not remote communities, in hotspots of linguistic diversity and poverty (Romaine 2014); rather, for the most part they may be considered primarily urban, somewhat privileged, communities.[2] This makes them no less significant: the phenomenon

2.   This paper focuses primarily on the urban community in Colombo. In earlier scholarship, the Colombo dialect, meant to represent SLM as a whole, is the one usually described, though

of urban linguistic diversity is receiving increased attention, in particular in this era of modern-day globalisation (Endangered Languages Alliance 2012; Siemund, Gogolin, Schulz & Davydova 2013; Lim 2013–2016). While the pragmatic needs of global urban cultures often trigger rapid language shift and loss, resulting in dramatic endangerment situations, it is often in the diaspora in urban contexts that ancestral language practices or cultural vitality are maintained.

2.   The days of *java* and *patois*

2.1   The Sri Lankan Malays

The Malays of Sri Lanka came to be through one of the central practices of Western colonialism, namely, the movement of subjects from one colonised region to another. In this way sizeable communities of people from Indonesia (the Dutch East Indies) and Malaya were settled on the island of Ceylon (present-day Sri Lanka) through various waves of deportation, the majority tracing their ancestry to the communities brought over during Dutch rule (1656–1796) and during British rule (1796–1948).[3] At least three different communities could be distinguished. First, there was a rather sophisticated diaspora of noblemen – nobility exiled during Dutch occupation of the East Indies – who typically would be deported together with their families, as were political exiles from different corners of the Indonesian archipelago and beyond, including Java, Borneo, the Moluku and Goa, among other places. Second, the largest group of people attributed a 'Malay' origin came as soldiers also from disparate places such as Bali, Java, Riau, Ambon and peninsular Malaysia, imported first by the Dutch to form a 'Malay' garrison,

---

variation between the different communities is briefly acknowledged (Saldin 2001); the issue of variation is addressed in more recent analyses (see e.g. Ansaldo, Lim & Nordhoff 2006). Two points bear mention here. First, while there is a strong sense of identity and separateness for each of the different communities (Ansaldo and Lim fieldnotes 2003–2006), they nonetheless all identify themselves as Sri Lankan Malays. This has surely been the case since colonial rule where this 'Malay' diaspora is attested as a close-knit community, in which contacts between the different Malay/Indonesian ethnicities as well as the different social extractions were maintained through the ranks of the army as well as through common religious practice (Ricklefs 1974). Second, however, is the significance of the distinction between the more urban and the more rural communities, where the degree of centralness as opposed to periphery has an impact on issues such as the degree of endangerment faced as well as the implications for citizenship participation (for an account see Ansaldo & Lim fc.).

3.   It is possible that the community based in the Slave Island district in Colombo may have been there during Portuguese rule (until 1656).

which would become the Ceylon Rifle Regiment under the British who continued the same practice (Ricklefs 1974). The soldiers too could also be accompanied by their wives, a practice encouraged during Dutch and British rule (Sourjah 2003; Ansaldo 2008, 2009). A third group comprising convicts, slaves and indentured labourers was surely present from as early as Portuguese occupation, and such importation continued through both Dutch and British rule. Contacts between the groups were indeed quite frequent, due among other reasons to the practice of employing noblemen as officers of the troops, master-servant relations and a common, Islamic faith (Hussainmiya 1987, 1990; Ansaldo 2008, 2009). Overall, the community referred to collectively as 'Sri Lankan Malays' constituted not just single individuals but also included family, retinue, and network ties, whose origins, and thus ethnic and linguistic backgrounds, are very heterogeneous, covering an area from Northern Malaysia to the easternmost provinces of Indonesia (Hussainmiya 1987, 1990). In an earlier era, in fact, these peoples were known as *Ja Minissu* by the Sinhalese and *Java Manusar* by the Tamils: 'people from Java' (Saldin 2003:3). It was the British who, upon finding a community who spoke 'Malay', attached the corresponding ethnic label to the group, and it is this designation 'Malay' that has persisted.

Since the late 1800s, the Malays have comprised approximately 0.33% of the population, and are still today a numerical minority in Sri Lanka, with the majority Sinhalese comprising two-thirds to three-quarters of the population, and a significant minority of Tamils comprising a quarter of the population. They are also a minority in name. They were subsumed in the Ceylon Citizenship Act of 1948 – together with the Sri Lankan Moors (Tamil-speaking people tracing their ancestry to Arab traders who arrived in Sri Lanka between the 8th and 15th centuries) and Indian Moors (from India) – as 'Moors' (Official Website of the Government of Sri Lanka) or 'Muslims' (Sri Lanka Government Web Portal). They have thus not had a distinct identity as 'Malays' at this official level. Only very recently are they mentioned in name, albeit grouped together with other communities as "Moors, Malays, Burghers (of Portuguese and Dutch descent) and others" (Government of Sri Lanka 2014).[4]

As a small minority in contact with the major ethnic groups in Ceylon, it is not surprising that, over the centuries, various aspects of the Malay / Indonesian community evolved to become 'Sri Lankan'. While their religious practices were maintained in the Muslim tradition, the community assimilated numerous cultural traditions of the two dominant ethnic groups of the island, the Sinhalese and the Tamils. For instance, until only very recently, the Sri Lankan Malay women

---

4. In censuses, though, Sinhalese, Sri Lanka Tamil, Indian Tamil, Sri Lanka Moor, Burgher, Malay, and Other are separate ethnicity categories.

have worn the South Asian *sari* as their traditional dress, rather than *baju kurong* or sarong *kebaya* as in Malaysia and Indonesia (Saldin 2003: 1), and weddings involve payment of a dowry as in Hindu practice, in contrast with Islamic tradition which only involves the groom's payment of *mahar* to the bride's father (Saldin 2003: 67).[5]

Linguistically, the early Malays would have been speakers of the Malay lingua franca that had existed since the 1st millennium AD in the monsoon Asia region, particularly along the trade routes between southern China and northwest India, most often referred to as Bazaar Malay (Adelaar & Prentice 1996), as in example (1). This Malay variety would have been in contact with two typologically distinct adstrates: colloquial Sinhala, the dominant language of the population of Sri Lanka, and Lankan Tamil, spoken by, amongst others, traders and plantation workers, illustrated in (2), (3). What evolved was a unique restructured variety now known as Sri Lanka Malay (SLM), a mixed language of trilingual base, with lexical items predominantly from Bazaar Malay and grammatical features, including V-final word order, number and case morphology, and agglutination, from Sinhala and Tamil (Aboh & Ansaldo 2007; Ansaldo 2008, 2009; Ansaldo & Nordhoff 2009), illustrated in (4), contrasted with Standard Malay in (5). SLM is usually considered a creole in the literature, with more recent work capturing the process of language creation as one of metatypy (Ansaldo 2009, 2011) leading to a hybrid profile of Lankan grammar and Malay-derived lexicon.

(1) *saya tak tahu cakap melayu*   (Trade) Malay (Austronesian)
     I   NEG know speak Malay
     'I can't speak Malay'

(2) *eyaate hungak salli   tiuna*   Sinhala (Indo-European)
     he-DAT much   money exist-PAST
     'he had a lot of money'

(3) *ongalukku ayare   teriyumaa?*   Tamil (Dravidian)
     you-DAT   he-ACC know
     'do you know him?'

(4) *samma anakpada manahari iskulnang   arpi*   Sri Lanka Malay
     all   child-PL   everyday school-DAT DUR-go
     'all the children go to school every day'

---

5.   Due to the religious affinity with Muslim Tamils, there have been historical and linguistic speculations suggesting that Sri Lankan Malay communities descended from Tamil-Malay intermarriages. This is however a mistaken view based on skewed interpretation of historical sources and not supported by recent historical and linguistic evidence (Ansaldo 2008).

(5)   *kanak-kanak semua sehari-hari pergi sekolah*   (Standard) Malay
      child-2   every one.day-2   go   school
      'all the children go to school every day'

In addition to their restructured vernacular, the Malays in Sri Lanka have also been noted for being the most multilingual of all the co-existing communities on the island – Sinhalese, Tamils, Burghers – having in their repertoire, alongside their SLM vernacular, the major languages spoken on the island, namely Sinhala and Tamil; in some strata, their repertoire also included the language of the colonial power (Lim & Ansaldo 2007; Ansaldo & Lim 2014).

Such a multilingual repertoire afforded many Sri Lankan Malays the opportunity to function in both Dutch and British Ceylon as intermediaries between colonisers and locals, as they were proficient in the languages needed to interact with all parties concerned. For instance, a majority of the ancestors of the Colombo, Kandy and Hambantota communities were Javanese nobility exiled during the wars of succession in Java during Dutch rule.[6] The proficiency in Dutch of the older Javanese allowed them to be appointed *Hoofd de Maha Badda* (Sinhala *maha badda* 'great trade', referring to the cinnamon industry first established by the Sinhala king in the 1500s for Portuguese trade) or *Hoofd de Cinnamon*, namely, the 'captain' supervising the cinnamon gardens, the spice being one of the most precious commodities during Dutch rule. With increased production of cinnamon, these superior officers would be rewarded with more power, promotions and privileges (Burah 2006). Most of the exiles became enlisted in the military, and were later retained under the British as members of the Malay Regiment (as well as in the Police and the Fire Brigade, COSLAM 2002), where, although they dropped their royal titles, they did nonetheless maintain their status as was the practice of the time (Burah 2006: 46–47). After the disbandment of the regiment in 1873, many of the Malays joined the tea estates and, with their proficiency in English, functioned as intermediaries between the English superintendents and the Indian labour force (Saldin 2003: 10). In short, as a consequence of their privileged origins, at least in some cases, and their multilingual repertoire which also included the colonial language, the Sri Lankan Malays have held a status amongst the communities that has been high.

Sri Lankan Malay communities are found around the island, in both the urban centres and the rural peripheries, as a consequence of their diverse origins and settlement patterns. The communities vary in their socioeconomic and

---

6.   Official documents of 1792, for example, list 176 individuals belonging to 23 families of royalty and nobility exiled together with their families from Java, Batavia and Sumatra to Ceylon (Burah 2006: 44).

Table 1. Sri Lankan Malay communities

Community	Characteristics
1. Colombo	Middle-upper class community in capital city; restricted usage of SLM in old-middle generations; common Sinhala (and some Tamil) competence; English fairly fluent to native speaker competence; standardising in Malay; no SLM in younger generation
2. Slave Island	Lower class community in a poor district of Colombo; strong Tamil influences; no English
3. Kandy & other Upcountry	Middle-lower/rural class communities in the central hill country area; SLM in old-middle generations, and in some younger generation; Sinhala competence; some English proficiency, especially in younger generation
4. Hambantota	Community on the south coast, traditionally heavy Sinhalese-speaking area; SLM in old-middle generations; often trilingual with Sinhala and Tamil; limited English
5. Kirinda	Fishing community on southeast coast; SLM dominant in all generations; fully trilingual with Sinhala and Tamil, especially in middle-younger generations; English limited to a few individuals

educational status, and their linguistic repertoire and communicative practices, as summarised in Table 1 (Lim & Ansaldo 2006, 2007, adapted from Ansaldo 2008). This holds implications for ethnolinguistic vitality, and language shift and endangerment. This paper focuses primarily on the urban community in Colombo.

In spite of their low symbolic capital within the modern nation state, their ethnolinguistic vitality (after Giles, Bourhis & Taylor 1977) has always been high. There has been much awareness and expression of their culture and ancestry (e.g. Saldin 2003; Burah 2006), and there are a large number of social and cultural groups, including, for example, the Sri Lanka Malay Confederation (SLAMAC) (the umbrella organisation), the Sri Lanka Malay Rupee Fund, the Conference of Sri Lanka Malays (COSLAM), and Malay Associations of the communities located around the island, which are all extremely active in the organisation of regular social, cultural, commemorative and fund-raising activities and initiatives (Ansaldo & Lim fieldnotes 2003–2007). Given the symbolic and social capital of the community as outlined above, and their dense and multiplex networks, it is not surprising that SLM was widely spoken as a home language for generations (Hussainmiya 1986).

2.2 The Peranakans

The Peranakans are descendants of southern Chinese seafaring traders who settled in the Malay archipelago from at least the 17th century and who married

non-Muslim natives of the region, such as Balinese or Batak slaves. Settling primarily in Malacca and Penang in peninsular Malaysia, and in Singapore – British colonies on the Strait of Malacca which were amalgamated in 1826 to form the Straits Settlements – the Peranakans comprised one of the earliest and largest groups of the influential class of Chinese capitalists in the region. In contrast to the Chinese who returned to China, the Peranakans, even with their trading movements between the ports of southern China and Southeast Asia, always returned to the Straits Settlements and considered Malacca and Singapore their home. Until as recently as the 1950s, only the Straits Chinese (as the Peranakans were also called) could be considered 'permanent', 'native', or indigenised Chinese communities in the region (Song 1923/1967, in Kwok 2000: 205). The new hybrid culture that emerged in this context shows unique traits that set the Peranakans apart from other Chinese, the more indigenous local populations, and other ethnically mixed groups (Tan 1988b; Rudolph 1998). Non-linguistic examples include a mixed *nyonya* cuisine consisting of Chinese culinary practices largely influenced by Malay traditions, and the wearing of Malay / Indonesian *sarong* and *kebaya*, instead of the Chinese dress, by the women. These contrast with the retention of Chinese rituals, such as religious practices mentioned above and traditional wedding customs involving imperial era wedding costumes (Tan 1988b: 299). According to some observers, the Peranakans had "lost touch with China in every respect, except that they continued to uphold Chinese customs, and to practice, in variously modified forms, the social and religious practices of the forefathers" (Lim 1917, cited in Kwok 2000: 202; Tan 1988a: 47). (See Lim 2010a and 2016 for detailed accounts of the Peranakan community.)

In the linguistic world, in particular in creole studies, they are probably best known for their vernacular, Baba Malay (BM), a restructured variety of Malay with substantial southern Chinese (primarily Hokkien) influence. As can be seen in examples (6) and (7), amongst other things, the BM pronominal system derives from Hokkien, and the word order too is Sinitic. Both the sociocultural identity of the Peranakans and their vernacular as a creole language have received scholarly attention (e.g. Tan 1988a, b; Pakir 1986; S. Lim 1988; Rudolph 1998; Ansaldo & Matthews 1999), with their being compared with the more typical creole communities (Ansaldo, Lim & Mufwene 2007).

(6) *Gua punya bilik*   Baba Malay
     1SG POSS   room
     'my room'

(7) *Bilik saya*   Malay
     room 1SG
     'my room'

Due in no small part to the fact that they had been in the region longer, more continuously and more permanently than the other Chinese immigrants, the Peranakans formed the larger proportion of the influential class of Chinese capitalists in the Straits Settlements, having established themselves in the mining of gold and tin, the large-scale commercial agriculture business (in gambier, pepper, tapioca, and especially rubber), import-export business, and other economic enterprises that had been drawing Chinese to Malacca for years (Tan 1988a: 48). By the time of the European exploitation colonisation of the region in the 19th century, most Babas of Malacca had accumulated much wealth and become prestigious sub-groups in the region, forming separate communities of their own. In particular they distinguished themselves from the later Chinese immigrants, referring to them derogatorily as *sinkeh* 'new guests', i.e. 'new arrivals', whom they considered poor and with low social status (Tan 1988a: 45). In Singapore as well, the Babas were a class apart from the other ethnic groups. Although small in number ('Malacca men' comprised only 2.5% of the Chinese population in 1848, growing to just 9.5% in 1881), their social and economic influence was disproportionately strong in comparison, and they formed an important sector of the local elite (Kwok 2000: 202–204). By the 1920s, Singapore-born Peranakans controlled the pineapple industry, and most of the rubber which was cultivated, at one time more than 8000 hectares in Singapore as well as in Malaya – which, together with tin, drove Singapore's prosperity in the late 19th and the 20th centuries (Liu 1999: 98). In Malacca, the well-off Baba were able to take over the houses of the great Dutch merchants in Heeren Street which then became "the fashionable and aristocratic resort of the Chinese" (Braddell 1853: 74). In Penang, it was also noted that the Chinese "who have long been settled in the place, and who have wedded native wives, dwell in large and elegant houses environed with fruit and flower-gardens" (Thomson 1875: 13). In Singapore, the Peranakans were wealthy enough to afford weekend retreats or second homes in the form of seaside bungalows – some with swimming enclosures – in the East Coast of the island, an increasingly attractive residential area from the end of the 19th century (Liu 1999: 148).

Just as in Ceylon, in the politics of segregation introduced by the Dutch in Southeast Asia (Reid 2000) and continued by the British, individuals of mixed origin were used as middlemen, merchants and interpreters between the colonial administration and local population and newer arrivals. Many of them worked for the British (as well as, in earlier days, the Dutch) East India Company (Tan 1988a: 51f.), and their command of the English language meant closer contact with British administrators and merchants (Nathan 1922: 77). Furthermore, their multilingual repertoire which comprised Baba Malay, Bazaar Malay, Hokkien and possibly one or two other Chinese languages, as well as English, and their knowledge of local ways afforded them a significant role as intermediaries between

Europeans, locals, and Asian newcomers (Tan 1988a; Kwok 2000; Lim 2016). All this together with their business acumen gave them predominance in the commercial sectors (also see Ho & Platt 1993: 8–9), and they were considered the best educated, wealthiest and most intelligent section of the Chinese community (Nathan 1922: 77).

English was already becoming an increasingly important language of Southeast Asia, especially British Malaya, from the early 19th century. Being wealthy merchants of high social standing, the Peranakans not only held a high regard for English-medium education but crucially were one of the earliest and privileged few in Singapore who had access to it, and sent their children, including girls – a rare occurrence in that era – to English-medium schools. By the early 1800s, members of the community had established four educational institutions in Malacca and Singapore which were especially important to the development of the community (Tan 1988a: 52). The establishment of the Queen's Scholarship in 1885 for British subjects in the Straits Settlements further enabled a few Peranakans to be educated in higher institutions in Britain, producing scholars and leaders (Tan 1988a: 65, 82). Already in earlier days the Peranakans were noted to have spoken English "tolerably well" (Earl 1837, in Tan 1988a: 50). By the mid-19th century their ability to converse in this colonial language had strengthened their prominent socio-economic position within other local communities in relation to the British, to the point where they were in fact sometimes referred to as the "King's Chinese" (Tan 1988a: 53), in relation to the King of England.

In other ways they realigned themselves culturally, distinguishing themselves from the continuously increasing population of China-born immigrants by their local (Malayan) orientation and their pro-British sentiments (Tan 1988a: 54f.). In their social clubs "to which they will admit no native of China … they play billiards, bowls, and other European games, and drink brandy and soda ad libitum" (ibid.). Not an uncommon observation then was for Peranakans "on being asked if they were Chinamen [to] bristle up and say in an offended tone 'I am not a Chinaman, I am a British subject'" (Vaughan 1879). Identifying politically with the British (Kwok 2000: 205), they formed the Straits Chinese British Association (SCBA) in August 1900,[7] with an admitted aim to promote trade with, and foster loyalty to, the British Empire (Song 1923: 319).

---

7.   The Malacca branch of SCBA was formed in October of the same year (1900), and the Penang branch was founded later in 1920. The Associations are all still extremely active to date, with the Singapore one renamed The Peranakan Association.

3.   The times they are a-changin'

Thus far, we have seen how, during the colonial period, in both communities, the language of the colonisers – earlier Dutch, then English for the Sri Lankan Malays; English for the Peranakans – was acquired and/or wielded to their advantage in attaining greater access to symbolic and economic capital and mobility. Long before what we consider today's 'globalisation', the English language in British colonial contexts was already a main player in language shift scenarios. Subsequently, the language policies instituted in both nations at independence in the mid-20th century, with ideologies of assimilation and/or internationalisation, consolidated the shift, as outlined below. (See Lim 2010b and 2013 for accounts of Singapore and Sri Lanka.)

3.1   Sri Lanka

While English education was introduced to the British subjects of Ceylon during British rule, as in all the other colonies, English was available only for a small and unbalanced proportion of the population. The Colebrooke-Cameron Report estimated that in 1828 less than 2% of the population were in school, and for those 250,000 under the age of puberty, only 800 were taught in the English language. Significantly, most of those who received English education were those in American mission schools in the Tamil north (Colebrooke 1831, cited in Bailey 1998: 210f.). At independence there were more missionary-built schools in the Tamil-dominated north (Jaffna) than in the rest of the island. Ceylon Tamils comprised only 12.4% of the population in 1946 (distinct from Indian Tamils, agricultural workers who were regarded as stateless persons but constituted another 10.4% of the population). Nonetheless, with their resource of English, they were, conversely, well represented in government service, as well as in medicine and law, far more than their share of the population (Bailey 1998: 216). Such a pattern continued through early post-independence Sri Lanka, and it is perhaps not surprising that, as a result, the dominating sentiment was of Tamil favouritism under colonial rule. Leading up to independence in 1948, with a prime minister who was extremely concerned about ethnic and religious harmony and who envisioned a multicultural, secular democracy and a multiracial state that did not favour any ethnicity or any section of any ethnicity, the Legislative Council in 1944 made Sinhala and Tamil the official languages of Sri Lanka. Later, with the rise of Sinhala nationalism, the new coalition government led by Bandaranaike in 1956 proclaimed, in the 'Sinhala Only' Act, Sinhala as the sole official language of Sri Lanka. This has been seen to have been catering to the large rural Sinhalese

electorate, allowing the Sinhalese scholars to enter political and economic domains, and rectifying past injustices, as English would no longer hamper Sinhalese economic and social development. In spite of the official line, there was a continuing emphasis on English as the language of administration; it remained an important language in the public sector, which was a major provider of employment (Fernando 1996).

A decrease in teaching and learning English after independence also occurred: in the 1950s, while English was not officially banned from education, it was widely believed that it was forbidden to have an English stream in schools and English-medium schools were considered illegal (Gunesekera 2005: 76–77). Ironically, English continued to be a valuable language, because it was the language of commerce, science, technology and a host of other functions in Sri Lanka and internationally. In particular, when the state decided to liberalise the economy in 1977, the public sector shrank with the privatisation of many state-owned companies. At this point, many Sri Lankans, who had received public education in Sinhala or Tamil, and thus had little or next to no English competence, had to face stiff competition for jobs in the private sector which required English, or now limited jobs in the public sector. Thus the children from the elite or rich and urban Sinhalese or Tamil families, who had received private education and had learnt English, had the advantage in employment in particular in the private sector. Nationalism had simply reinforced elitism and exclusivism for English in Sri Lanka.

The implications of such language policy in concrete terms meant that, in the urban Colombo community, where the level of English-medium education was already high, Sri Lankan Malay parents and grandparents made the conscious decision to speak to their children in English in the home domain (Saldin 2001, 2003; Lim & Ansaldo 2006, 2007), in order to provide them a resource recognised as requisite for communication and advancement internationally – "the key to a good job and a comfortable life" (Saldin 2003: 76). The general pattern displayed is a clear shift to English from SLM in the home domain. As a result, by the end of the 20th century, the community was typically showing strong linguistic vitality in SLM in the oldest to middle generations and rapidly decreasing linguistic competence to nil in the vernacular in the young generation (Ansaldo & Lim fieldnotes 2003–2006; Lim & Ansaldo 2006, 2007). SLM now had a mere fifth position in the community, after Sinhala, Tamil, English and Arabic (the last in the religious domain). SLM in the urban community is no longer a home language for the younger generation of Sri Lankan Malays.[8]

---

8. Such a shift is found primarily in the urban, English-educated Colombo community, and to a lesser extent in more recent years in the other large urban centres such as Kandy (Sebastian Nordhoff p.c. 2007; Ansaldo & Lim fc.).

We can see the explanation for this choice in terms of the capital the various languages possess in the country. In the local linguistic market – of school, profession, politics – Sinhala is recognised as a necessary capital and accepted without battle; it has in any case always been in the Sri Lankan Malays' repertoire. Similarly, English has been an important variety in their linguistic repertoire; it was a language which allowed the Sri Lankan Malays many privileges as colonial subjects, and in postcolonial Sri Lanka a multilingualism including English was recognised as crucial to the Sri Lankan Malays. In consequence, SLM which has a low capital in the local linguistic market was forfeited, and thus became not only a minority and marginalised language, but also an endangered one.

In the early years of the 21st century, the SLM community recognised the shift that was occurring and the potential obsolescence of their variety, and took steps towards its revitalisation. However the variety selected for revitalisation activities was Standard Malay of Malaysia, due to several factors (detailed in Lim & Ansaldo 2007; Ansaldo & Lim fc.). First, because previous publications on SLM classified it as a creole, the community's perception of their own language was less positive, viewing it as an 'imperfect' code and an ungrammatical dialect of Malay (e.g. Thaliph 2003; Colombo SLM community p.c. August 2006). Second, in contrast with the status the community has in their own country, greater recognition was forthcoming instead from Malaysia (and also Indonesia), in at least two significant and related thrusts, both clearly seen as arising from scholarly and transnational contexts which have associated symbolic and material markets. One of the objectives of Malaysia's Institute of Malay Language and Culture has been "to get in touch with Malays in different parts of the world and teach them the real Malay" (T.K. Azoor p.c. January 2006): one of the realisations of this was the organisation in Colombo of language classes in the Standard Malay (StdM) of Malaysia (*Bahasa Melayu*). Moreover, the Malaysian High Commission in Sri Lanka in those years demonstrated interest in, and strong support for, the Sri Lankan Malay community, and provided aid in terms of student scholarships for undergraduate and postgraduate studies in Malaysia, as well as in job market openings, with one of the requirements of the latter being competence in basic Malay. It is not surprising then that it was with Malaysia that the urban Sri Lankan Malay community aligned themselves then, both in terms of language and identity. After two pioneering courses in StdM in 2002, eight of the best students – note: from the Sri Lankan Malay community – underwent a teacher's training course in Malaysia and then conducted regular classes in StdM for the community. During that period, in the annual *Hari Bahasa Melayu* (Malay Language Day) organised by the community in Colombo in August 2006 and 2007, activities such as essay-writing and oratory contests were conducted for both SLM and StdM (Ansaldo & Lim fieldnotes 2006–2007; Lim & Ansaldo 2007). Ironically, when

StdM-speaking Colombo Malays attempted to communicate in Malay with Sri Lankan Malays from the other communities, they were not intelligible to each other. In short, the promotion of StdM – far from revitalising SLM – was further endangering the vitality of SLM in the urban environment.[9]

3.2 Singapore

Leading up to Singapore's independence, English was made one of the four official languages – the 'neutral' language not linked to any ethnic group, and associated with modernity and progress – with Malay, Mandarin and Tamil the official 'mother tongues' each associated with the three official races of Malay, Chinese and Indian. English became the medium of instruction in all schools by 1987, and Singapore became for all intents and purposes an English-dominant society. With the Peranakans categorised as 'Chinese' in Singapore's official racial classifications, young Peranakans studied Mandarin as a second language in school rather than have Malay reinforced. Singapore's modern-day ecology also afforded young Peranakans much less exposure to BM: the shift from BM to English was already in place in the community, and other factors, such as the change from

---

9. In sharp contrast are the more rural SLM communities such as that found in Kirinda, which is relatively peripheral as a small fishing village on the southeast coast. 90% of the village are Malay, and they comprise some 4% of the 46,000 SLM population. With a dense and multiplex social network, limited educational and employment opportunities, they exhibit strong maintenance of the vernacular across all generations. The variety of SLM in Kirinda is structurally distinct from other SLM varieties, where, lexically and grammatically, there seems to be a stronger influence of Sinhala and Tamil; trilingualism in SLM, Sinhala and Tamil is very common in all generations (Ansaldo 2008). There was thus no endangerment of the language here; on the contrary, the language displays a high vitality, as the dominant language of all generations of the community, spoken in all domains, even as the working language in the Tamil-medium *madrasah* (school). Kirinda is often said to be the only fully vital community of Sri Lankan Malays in which a young generation of speakers of a SLM variety as first language can be found; similar vitality has also been documented in the Upcountry communities (Sebastian Nordhoff p.c. Jan 2006). Ironically, the 'revitalisation' efforts in Colombo and increased prominence of StdM in the discourse on language led to the Kirinda community becoming even more explicitly aware of the more prestigious variety and the possibilities it might – in theory – hold for them. In 2006, plans were underway for StdM to be taught in the village school as a subject, to children who in fact were native speakers of SLM, as well as to be used as a default written language, for example, in the signs (e.g. 'no shouting'; 'show respect') displayed around the school (Lim & Ansaldo 2007). In other words, while SLM was never endangered in Kirinda, the discourse of the urban centre – involving the promise of increased economic mobility via the 'revitalisation' of a standard variety – was impacting on the periphery, and threatening to displace their vernacular.

extended to nuclear family units, meant that BM speakers of the grandparents' generation were not as present in daily domains as in the past. Furthermore, the demise of BM-speaking generations meant an overall reduction of BM (speakers) in the ecology (Lim 2010a). In the mid-20th century, even while most Peranakans still identified themselves with BM (Tan 1988b) at least symbolically, English grew in importance in the Peranakan community as a lingua franca, resulting in an increasing shift to English as the only vernacular and identity marker (paralleling the broader situation in Singapore as a whole). While BM had competed fairly well with English as a lingua franca until the early 20th century, by the end of the 1960s English had almost completely prevailed de facto as the means of interethnic communication par excellence (Rudolph 1998: 335). Baba Malay is now considered an endangered language, classified as vulnerable (Alliance for Linguistic Diversity, n.d.).

4.   The art of losing

To return to the trope in this paper, "[t]he art of losing", to repeat poet Elizabeth Bishop, "isn't hard to master / so many things seem filled with the intent / to be lost that their loss is no disaster". The series of losses presented by the poet – including her mother's watch, houses, cities, continents – come with reassurances that all may be absorbed: one might "miss them, but it wasn't a disaster". In this section I explore the question of whether ancestral language loss in a community, such as that found for SLM and BM, is a disaster – in terms of the irrecoverable loss of linguistic and cultural diversity, including their correlation to biodiversity and sustainability, as well as the implications for cultural identity and linguistic human rights – all of which are widely recognised by scholars as well as supranational bodies such as UNESCO. I then continue with a consideration of whether such 'loss' is a development that may in fact be embraced.

4.1   Value for linguistic science

As with most other languages considered endangered, the imminent obsolescence of these varieties together with their value for linguistic science have been recognised by scholars. I briefly mention some aspects below.

SLM, together with a very few other varieties of the region (e.g. BM, Cocos Malay) – unlike its better-known Caribbean 'creole' counterparts – is noted (Ansaldo 2008, 2009; Ansaldo & Lim fc.) for being typologically in a unique

position of providing us with an ecology in which no Standard Average European[10] variety is involved in the dynamics of contact. Furthermore, with Sinhala and Tamil as its adstrates, the languages involved in the formation of SLM varieties come from three distinct language families, Austronesian, Dravidian and Indo-European, with marked typological differences amongst the languages in contact. As such, SLM can help shed light on issues of universality and specificity in contact-induced language change (Ansaldo 2008, 2009). The relationship between the three language communities involved in its evolution – that is, the Malays, Sinhalese and Tamils – was also of an altogether different type than the better-known creole scenarios of exploitation, slavery or intermarriage between coloniser and slave,[11] and thus provides contrastive material for our understanding of typologies of language contact.

BM, in addition to being valued as a contact language, has also recently been recognised for its significance in the evolution of New Englishes (Lim 2011, 2014, 2016). The pattern of word- or phrase-final prominence found in Singapore English (SgE) contrasts with all other contact varieties in which tone has evolved, such as Nigerian English or Hong Kong English, where the general pattern locates high (H) tones on what would be stressed or accented syllables at word- or phrase-level. SgE's unusual prosody can be explained if we consider the Founder Principle in the ecology paradigm (Mufwene 2001, 2008), which suggests that the founder population in an ecology exerts a strong influence on features, an influence which persists in the emergent variety, and if we consider the Peranakans as a founder population in Singapore's ecology, one with significant economic and social prominence in the ecology. Word-/phrase-final prominence which is widespread in many Malay varieties, is documented in BM, and also developed in PerE as a result of contact in the Peranakans' multilingual repertoire. The Peranakans, though a small minority, were clearly dominant in the external ecology, as outlined earlier, due to their political, economic and social status, and their position as intermediaries, and later as teachers. As early English adopters, crucially during the British colonial period, theirs would have been the early features influencing the emerging variety of SgE.

In short, SLM and BM are significant varieties for informing studies of language contact, language evolution as well as cultural creolisation.

---

10. 'Standard Average European' is a term introduced by Whorf (1941/1956) to refer to Indo-European languages of western Europe on the basis of their sharing structural similarities. More recently, Haspelmath (2001) uses the term to refer to a number of traits characterising a Sprachbund, or linguistic area, defined by a core of twelve features that can be considered 'euroversals'.

11. Though the type of colonisation and the evolution of the contact variety are comparable.

4.2   Capital in languages of wider communication

A positive note for the fields of creole studies and language endangerment is that both SLM and BM have by now received substantial documentation. A related but separate issue is whether the varieties should see revitalisation in their communities, for the sake of cultural identity and ethnolinguistic vitality of the community. Scholars certainly have held that "there is no language for which nothing at all can be done" (Fishman 1991: 12), and programmes for reversing language shift were already being outlined two decades ago (e.g. Fishman 1991). Revitalisation activities are on the schedule in numerous projects (see Hinton 2011).

Other scholars have however queried the attention and activities involved in such an endeavour of halting language shift and/ or revitalising endangered languages. The late Peter Ladefoged suggested that, since language death is a natural part of the process of human cultural evolution, linguists should simply document and describe languages scientifically, but not interfere with the process of loss and revitalisation. A poignant anecdote, often cited but well worth mentioning, provided by Ladefoged (1992: 810–811), is as follows:

> Last summer I was working on Dahalo, a rapidly dying Cushitic language, spoken by a few hundred people in a rural district of Kenya. I asked one of our consultants whether his teen-aged sons spoke Dahalo. 'No,' he said. 'They can still hear it, but they cannot speak it. They speak only Swahili.' He was smiling when he said it, and did not seem to regret it. He was proud that his sons had been to school, and knew things that he did not. Who am I to say that he was wrong?

In the urban Sri Lankan Malay and Peranakan communities, a shift to a regionally or globally more dominant language has in fact afforded significant opportunities for upward mobility. As mentioned in Section 2, having a language such as Dutch (in Dutch Ceylon) or English (in British India or Malaya) afforded the communities significant income-earning opportunities and positions as intermediaries. In the modern nation state, having languages such as English or Standard Malay in their multilingual repertoire affords the communities clear advantages in terms of economic capital in both the local and larger linguistic market.

In the case of the Sri Lankan Malays, this has been couched in terms of identity alignment (Lim & Ansaldo 2007; Ansaldo 2009; Ansaldo & Lim 2014), where the community is seen to preserve and represent their subject position by (i) not contesting their imposed identity (which is not negotiable) of 'Muslim' in the context of the nation-state; (ii) still maintaining their presumed ethnic identity as 'Sri Lanka Malay'; while (iii) aligning themselves with an assumed global Malay identity, one which is accepted and not negotiated, via the acquisition of Standard Malay. The new languages shifted to are accepted by the community as legitimate,

and valued as resources and not threats. Speakers have agency to opt for a 'shift' in which the economic and cultural functions of the language come together, and through such a choice the group is provided access to better education and enhanced political self-representation, while maintaining their subject position.

With the Peranakan community in the early 21st century, such a shift is no longer even viewed as a problem. As is evident in example (7), from the editorial of the Peranakan Association's (July/September 2002) newsletter, the actual code used by the community is not the issue, and the Peranakan identity and identification are clearly and positively recognised, whatever language is used, "be it [in] English, Malay or Chinese" as the editor writes.

(7)   We Peranakans have our own way or style of speaking that has become our trademark, which those outside the community recognise immediately, be it in English, Malay or Chinese. One Nyonya, for instance, tells me she is never surprised when people she meets for the first time straightaway say 'Ah, you are Peranakan, right?

This description explicitly reinforces the observation made earlier in this paper: that it has been and still is the multilingual repertoire of the community that has been significant for their existence and cultural identity, from the colonial era to the present day, rather than any specific language that is critical to their vitality, whether ancestral or emergent (Ansaldo & Lim 2014; Lim 2016).

4.3   Emergence of new varieties

Even if we accept that shift to a language of wider communication might bring increased economic capital, as discussed in the previous section, it is still often hard to counter the argument that the loss of a language is usually accompanied by the diminishing of cultural diversity and loss of intangible cultural heritage. Yet, this is not always the case. In some situations of language shift, the language of wider communication – such as English – with which the vernacular or ancestral language comes into contact, is, as noted by Woodbury (2005), at times adapted ideologically, if not always structurally, to communicative ends which are continuous with those earlier fulfilled by the ancestral language. This is observed, for example, in Aboriginal communities in southeast Queensland, Australia (Eades 1988: 97, 101):

While many Aboriginal people [in southeast Queensland] may speak English as their first language, the context of conversation has significant Aboriginal cultural and social aspects which lead to distinctively Aboriginal interpretations and meanings. While the chosen language code is frequently English, there are

important continuities in the ways language is used. … The Aboriginal priority on developing, maintaining, and strengthening social relationships is both reflected in, and created by, the way people speak to each other, whether the language variety is English, Aboriginal English, or Lingo [any Aboriginal language].

Similarly, in Koyukon communities in Alaska's interior, even in the rapid shift from Koyukon to English, cultural patterns are transferred (Kwachka 1992: 70–71):

The Koyukon people have been able to transfer and permute a very important cultural pattern at the discourse level, the tradition of narrative. … Although [stories from a distant time] are rarely told today, the narrative, as a social and rhetorical structure, has not only persisted but flourished.

In the Peranakan community in Singapore, the emergent variety is a distinct contact variety of English, Peranakan English (PerE), first documented in Lim (2010a). In written mode (based on a corpus of newsletters published by the Peranakan Association (1994–2008), PerE involves numerous lexical items of Baba Malay and Hokkien origin, conveying cultural practices, food, naming and address practices, exclamations, greetings, wishes and thanks, a sample of which are shown in the examples in (8).

(8) a. *dondang sayang* [Malay]: love ballad, originating in Malacca in the 15th century, influenced by traditional Portuguese folk music, now a traditional form of entertainment for Malays and Peranakans involving violin two Malay *rebana* 'drums' and a *tetawak* 'gong', in which singers exchange Malay *pantun* 'poetry' in lighthearted and sometimes humourous style.

   b. *biji saga seeds* [Malay *biji* 'seed']: small, hard, bright red seeds from the fruit pods of the red sandalwood, also known as coral tree or saga tree, a deciduous tree found in tropical and subtropical regions of the world; the seeds are used as beads in jewellery, leis and rosaries; they were also used in ancient India for weighing gold; the word *saga* is traced to the Arabic term for 'goldsmith'

   c. *kimpoh choh* [Hokkien]: maternal great grandaunt

   d. *May we extend to all readers a Selamat Tahun Baru* [Malay 'happy new year'] *and may you all enjoy panjang panjang umor* [Malay 'very long life'] *in the year of the Goat*

   e. *The Main Wayang Company would like to say a big KAMSIAH* [Hokkien 'thank you'])

In spoken form, PerE is even more clearly a single English-*Baba Malay* code (Lim 2010a), illustrated in example (9) (Lim 2010a: 336).[12] The speaker, an 81-year-old female, was typical of her generation in having spoken Baba Malay and English with her parents and later her husband, primarily English and Baba Malay with her siblings, and also having Cantonese and Hokkien in her repertoire. She studied in an English-medium school until she was 16, completed her Senior Cambridge examination and teachers' training college, and thereafter went on to become a teacher.

> (9)   Like drugs you know when you're under drugs … The babies come out crying crying. *Apa dia mo? mo?* Drugs. *Nanti* alcohol the same. They get into their system. *Kita semua tak drin[k]* … Keep yourself clean and healthy. Don[t] drin[k] don[t] drin[k].
> 'Like drugs, you know, when you're under the influence of drugs … The babies are born crying. What do they [the babies] want? Drugs. Then it's the same with alcohol – it gets into your system. We all didn't drink. … Keep yourself clean and healthy. Don't drink.'

It is clear that aspects of the ancestral culture involving address practices, food, certain cultural and religious practices, and terms of emotive import and value judgement are still transmitted in the emergent contact variety. In other words, the evolution of a mixed code affords the maintenance of culture even if the ancestral language is no longer maintained. Cultural diversity can be and is still maintained, and indexed, in an emergent variety. It may not be felt by older heritage language speakers to convey with precisely the same meanings and affect what is conveyed in the heritage language. But the emergent variety has the advantage of being a variety that is adapted to the immediate, current ecology, and a variety that is acceptable – native – to the younger generation: this means that it is in a good position, as seen in the Koyukon situation above, not only to survive, but to flourish.

4.4   Postvernacular vitality

Minority and endangered language communities such as the Sri Lankan Malays and the Peranakans were groups which saw their formation and evolution in the particular sociohistorical context of the Dutch and/ or British colonial eras.

---

12. The data presented in Lim (2010a) are more English-dominant, since PerE is the focus; in interactions between Peranakans in whom BM is more active, usually in the older generations, a higher frequency of BM is documented (Lim fieldnotes 2010–2015).

Writing about the British Straits Settlements, Bloom (1986: 360) recognises this very fact, pinpointing "the amalgam of Asian cultural traits and the English language in groups such as the Straits Chinese, Anglo Indian groups and Portuguese Eurasians, in particular in the Straits Chinese, [was] unique to the Straits Settlements, which made them an indigenous culture in a palpable sense". The ecology in which they were formed has however certainly changed – to the modern nation-states of Sri Lanka and Singapore respectively. A changing ecology, as already seen, often catalyses shift from and loss of an ancestral language. In this section, I explore in some detail the question of whether the overall cultural vitality of such communities is negatively impacted as well.

Indeed, for a period, the Peranakan community seemed not to have been evolving along with the times: as noted by Peter Lee, a core figure in the community, the community had chosen what has been seen in retrospect a self-imposed exclusivity during the 1960s–1970s, and remained in "ultra-conservative mode, lost touch with the reality of the world then and painted itself into a corner" (Yap 2008). The imminent obsolescence was predicted not only for their language but just as much for the cultural group itself, with predictions being made of "the dying out of the Peranakans" (Kwan-Terry 2000: 96). The new century has however witnessed a renewed vitality of both the Peranakan community in Singapore (Lim 2010a, 2014, 2016), and the Malay community in Sri Lanka (Ansaldo & Lim 2014, fc.), one that has complex layers.

On the one hand, in both communities, there is still symbolic maintenance or revitalisation of the vernacular and traditional culture. Within the Sri Lankan Malay context, while there already were publications by members of the community on their identity and language (e.g. Saldin 2001, 2003), as well as books comparing SLM with StdM (Saldin 2000; Thaliph 2003), further developments were largely the result of a large DOBES[13] documentation project which took place from 2004 to 2009. As outlined in Ansaldo and Lim (2014), documentation brought recognition to SLM as a variety with scientific capital, and thus the bias towards its creole or mixed nature was partly reversed. It became also obvious to the community over time that acquiring StdM was not the same as revitalising SLM, and eventually the latter became a priority in the community and affected the appeal of the former. Also as a consequence of the documentation project, SLM gained attention in the international community, and a number of academic publications reached the community – including a collaboration between academia and community in the form of an SLM dictionary (Saldin & Lim 2007).

---

13. The DOBES programme (Dokumentation bedrohter Sprachen) was an initiative started in 2000 by the Volkswagen Foundation in order to document languages that are potentially in danger of becoming extinct within a few years' time; see http://dobes.mpi.nl/.

All this strengthened the perception that the Sri Lankan Malay identity was a unique one, separate from that of Malaysia and Indonesia and unlike any other within Sri Lanka. The unprecedented attention that SLM started and is still receiving now, not only from Western intellectual powerhouses, but crucially from local linguistic institutions, has further strengthened the renewed prestige of a Sri Lankan Malay identity centred in, though not exclusively limited to, its ancestral language. While English is still the dominant language of the urban Sri Lankan Malay community in Colombo, members now do make a point of holding their meetings (at least partly) in SLM. Through a reevaluation of their linguistic capital, SLM communities are now focusing a lot of energy on symbolic maintenance and revitalisation of the ancestral language.

The Peranakan community in Singapore experienced an exceptional surge in cultural vitality towards the end of the first decade of the 21st century. This was due to a combination of several factors. A new group of leaders were elected to the community's association, who, as relatively prominent individuals in Singapore, garnered significant state support, both economic and symbolic, as well as presence in local media. Together these led to a level of cultural vitality in the community not seen in modern times (Lim 2014, 2016). There is documentation and maintenance of BM: works such as a BM dictionary and a collection of BM idioms (Gwee 1993, 2006) have been published; plays are written and performed regularly in BM; and a church in a traditional Peranakan district has been holding services in BM (though the BM-speaking priest passed away on 1 June 2013). The Peranakan Association's youth group has regular gatherings which involve not only engaging in Peranakan cultural activities but also revitalising BM by using it in popular culture, for example, in rap and hiphop by a group who set out to modernise Peranakan entertainment as a way to keep the heritage alive through innovation and to reach out to as many people as possible. Increased formal institutional recognition as an important cultural group and/or support by the state also resulted, for example, in a dedicated Peranakan Museum which primarily features traditional artefacts and practices, and the restoration of a traditional ancestral home of a Peranakan family as a heritage house, which opened in 2008 and 2009 respectively.

Both the Sri Lankan Malay and Peranakan situations as described thus far are best understood using the notion of postvernacularity (Shandler 2006), which describes situations where a language serves the purpose of identity-building within a community even after it has ceased to be used as a vernacular for daily communication, and has been receiving attention by several scholars in recent years, for example, with Yiddish in the US (Shandler 2006), Low German in Northern Germany (Reershemius 2009), and Breton (Hornsby & Vigers 2013). For the most part, studies on postvernacularity document the use of the postvernacular language in a

number of cultural practices, such as amateur theatre, music and folklore, translation, attempts to learn the language in evening classes, and in its primarily symbolic value and the tendency to preserve only the language's most colourful or evocative elements.

However both situations would appear to go beyond such postvernacularity. This is primarily due to the recognition by the communities of no longer needing to be constrained by the ideology of purity. The Peranakan community, for example, is explicitly embracing evolution in the 21st century, as is evident in a number of different initiatives, as follows. Even in more 'traditional' showcases of the community's culture, such as in museums, there is innovation. As part of the Sarong Kebaya Exhibition in 2011 at the Peranakan Museum, curators provided for an opportunity to 'do' Peranakan – notably in languages other than BM, for example, in their storytelling event, conducted in English. An exhibition on Peranakan artefacts including furniture, beaded and embroidered textiles and porcelain, curated by Singapore's Asian Civilisations Museum, which opened in 2010 at the musée du quai Branly in Paris was entitled *BabaBling* – using in its title the African American Vernacular English hiphop (though arguably now mainstream) term *bling* in a display of crossing. In the era of Web 2.0, both the Sri Lanka Malay Association and the Singapore Peranakan Association have websites and Facebook pages, which, apart from cultural items and symbolic uses of the language, are in English (though comments on the SLMA Facebook page may be in Malay or Sinhala). The Peranakan Association's stated aim is "preserving and promoting Peranakan culture" – with no mention of BM; and most of the contributions on the Facebook page *We Facebook in Nonya-Baba Peranakan Patois* – which "seeks to chat in Peranakan patois" – is in English. There is even a Peranakan presence in the virtual world of Second Life (Lim 2010a, 2014, 2016).

Two aspects merit highlighting. First, in both communities, the linguistic and cultural vitality the communities are experiencing is really embodied through all their language practices – that is, not only in their postvernacular languages of SLM and BM respectively, but also in their emergent languages of Sri Lankan English and PerE and SgE (which also include the mixing with the other languages in their ecologies), as already pre-empted in Section 4.5. Second, this challenges traditional wisdom in endangered languages literature of the link between language and identity and vitality of a community, in terms of the language being an essential part of a community's cultural identity and heritage. What is observed – in particular in the Peranakan community in Singapore – is an even stronger cultural vitality than before, in their emergent language.

But perhaps the most significant factor in their evolution has involved the relationship between the Peranakans as periphery and mainstream Singapore. This involved the appearance of Peranakan culture in mainstream media, in the form

of two locally produced television series which aired on local televisions channels in 2009. One of them, *Sayang Sayang*, was a sitcom for the local English television channel, and the other, *The Little Nyonya*, was a drama series produced for the local Chinese television channel. The fact that *The Little Nyonya* series was in Mandarin – one of Singapore's four official languages and the major lingua franca amongst the Chinese community since the mid-20th century – raised a furore in the community: why were Peranakans depicted as speaking Mandarin? However, such programming actually meant that the series – and the entire existence of the Peranakans – reached three-quarters of Singapore's population, viz. the Chinese-speaking majority, many of whom became acquainted with the culture for the first time. This led, almost overnight, to the appreciation of the culture outside the Peranakan community, with non-Peranakans in Singapore embracing and consuming its cuisine and material culture. *The Little Nyonya* marked a milestone in terms of the Peranakan community moving into the mainstream.

Once in the mainstream, the value of the Peranakan heritage came to be recognised, and Peranakan culture came to be considered as representing Singapore culture. While they were a previously marginal community in a territory (e.g. subsumed, as they are still, under the category 'Chinese' in Singapore), the Peranakans were proclaimed by Singapore's Arts and Information Minister Lui Tuck Yew in 2010 "multiracial emblems of [Singapore's] social mix". Peranakan culture and performances are now showcased on the global stage as the essence of being Singaporean: Peranakan culture and performances have been used to represent Singapore at international events, such as at the Asia-Pacific Economic Committee (APEC) forum in 2009, and the World Expo in Shanghai, China, in 2010 which featured Peranakan culture in the Singapore pavilion. Singapore's National Heritage Board's newest range of museum merchandise which "celebrates Singapore and what makes us unique" includes the use of designs from Peranakan ceramic tiles (commonly used in Singapore in the past to accentuate the architectural design of shophouses) in prints "to celebrate the intricate beauty and the unique identity of the Peranakan heritage".

5.   The art of evolving

To conclude, I return to the elements in the title and the quote at the start of the paper. Are Peranakans and Sri Lankan Malays trapped by the prediction in the traditional proverb *Dah sa chupak tak boley sa gantang* 'a quart will never make a gallon', that is, you cannot change destiny?

In the Sri Lankan Malays and the Peranakans, we see two communities who came into being through the practices of colonialism – 'globalisation' of an earlier

age. We can view their creation and the creation of their original vernacular as a logical and natural outcome of (pre-)colonial contexts in which a community negotiates between adequating for political and/or economic advantage (or lack of disadvantage) and retaining or manifesting some of its uniqueness, in the form of a 'creole' identification. Then, in the ecology in which they existed, the acquisition of English and/or some other language of wider communication, such as Standard Malay – and note, that this is the addition of a new language to an already multilingual repertoire – is simply part of that story. It comprises a resource that they found themselves in a position to acquire, which afforded them significant positioning. In other words, 'shift' is simply a continuation of the same trend, one which may be considered positive in terms of speakers having agency in constructing their linguistic identity via identity alignment, and where their multilingual repertoire and/or multicultural identity does not change, even if individual languages in their repertoire change.

Speaking of creole communities such as the Sri Lankan Malays, the Peranakans, as well as the Macanese of Macau, Ansaldo (2010) suggests that such recurring linguistic alignment is particularly salient in communities characterised by linguistic and cultural admixture; this is a consequence of their multilingual practices, which allow them a wide spectrum of linguistic negotiation, as well as a historical heritage of their sociolinguistic conditions of minority-migrant groups adept at constantly adjusting to changing environments. More generally, it is a point that has been made by Mufwene (2004, 2007, 2008) that the emergence of new languages through contact language formation should be viewed as a continuous ongoing process in the cycle of language birth and death, and one needs to consider not just the cost but also the benefits involved in language loss (Mufwene this volume).

It has previously been suggested that endangered contact languages are doubly marginalised (Garrett 2006: 177–178): marginalised amongst the world's languages in general, and then marginalised again amongst endangered languages. In contrast, in the new global order, in particular with multilingual communities, it has been argued that "late capitalism has shifted the positioning of the multilingual [and/or multicultural] periphery … As zones of both authenticity and multilingualism, former peripheries have much to offer" (Heller 2013). In the situations of the Sri Lankan Malays and the Peranakans presented in this paper, I suggest that, if a community is open to evolution in a changing ecology, the loss of their ancestral language need not be a disaster, and the loss of cultural vitality need not be an inevitability. The addition of a language of wider communication to their already multilingual repertoire is not a loss, but a gain in terms of capital and mobility, with the community's culture and identity not being limited to any one language. The emergence of a contact language affords the continued

embodying of substantial elements of culture and identity. And the repositioning of a periphery in the new global economy – as in the case of the Peranakans in Singapore, in particular – can transform a community to becoming an important site of authenticity, a source of added value, who does indeed have much to offer. The endgame is not just to participate in one's own postvernacular language community; rather, such evolution brings greater accessibility to and participation in the Centre, increased inclusion of the Other, and better adaptation for survival.

It is not accidental that the Singapore Peranakan Association's logo is a phoenix. As the mythical long-living bird that is cyclically regenerated or reborn, it is an apt symbol of the renewal that communities can undergo in order to survive – even thrive – in the changing ecologies that they experience, whichever language or languages they may speak. The art of losing is really the art of evolution and renewal.

References

Aboh, E.O. & Ansaldo, U. 2007. The role of typology in language creation: A descriptive take. In *Deconstructing Creole* [Typological Studies in Language 73], U. Ansaldo, S. Matthews & L. Lim (eds), 39–66. Amsterdam: John Benjamins. doi:10.1075/tsl.73.05abo

Adelaar, K.A. & Prentice, D.J. 1996. Malay: Its history, role and spread. In *Atlas of Languages of Intercultural Communication in the Pacific, Asia and the Americas*, S.A. Wurm, P. Mühlhäusler & D.T. Tryon (eds), 673–693. Berlin: Mouton de Gruyter.

Alliance for Linguistic Diversity. Baba Malay. *Endangered Languages*. <http://www.endangeredlanguages.com/lang/mbf>

Ansaldo, U. 2008. Revisiting Sri Lanka Malay. In *Lessons from Documented Endangered Languages* [Typological Studies in Language 78], A. Dwyer, D. Harrison & D. Rood (eds), Amsterdam: John Benjamins. doi:10.1075/tsl.78.02ans

Ansaldo, U. 2009. *Contact Languages: Ecology and Evolution in Asia*. Cambridge: CUP. doi:10.1017/CBO9780511642203

Ansaldo, U. 2010. Identity alignment and language creation in multilingual communities. *The Native Speaker and the Mother Tongue. Language Sciences* 32(6): 615–623.

Ansaldo, U. 2011. Sri Lanka Malay and its Lankan adstrates. In *Creoles, Their Substrates and Language Typology* [Typological Studies in Language 95], C. Lefebvre (ed.), 367–382. Amsterdam: John Benjamins. doi:10.1075/tsl.95.21ans

Ansaldo, U. & Lim, L. 2014. The lifecycle of Sri Lanka Malay. *Language Documentation and Conservation* 7: 100–113.

Ansaldo, U. & Lim, L. Forthcoming. Citizenship theory and fieldwork practice in Sri Lanka Malay communities. In *The Multilingual Citizen: Towards a Politics of Language for Agency and Change*, L. Lim, C. Stroud & L. Wee (eds). Clevedon: Multilingual Matters.

Ansaldo, U., Lim, L. & Mufwene, S.S. 2007. The sociolinguistic history of the Peranakans: What it tells us about 'creolization'. In *Deconstructing Creole* [Typological Studies in Language 73], U. Ansaldo, S. Matthews & L. Lim (eds), 203–226. Amsterdam: John Benjamins. doi:10.1075/tsl.73.11ans

Ansaldo, U., Lim, L. & Nordhoff, S. 2006. Variation and shift in Sri Lanka Malay: Preliminary reflections. DoBeS (Documentation of Endangered Languages) Workshop 2006. *MPI Nijmegen*. 15–16 June.

Ansaldo, U. & Matthews, S.J. 1999. The Minnan substrate and creolization in Baba Malay. *Journal of Chinese Linguistics* 27(1): 38–68.

Ansaldo, U. & Nordhoff, S. 2009. Complexity and the age of languages. In *Complex Processes in New Languages*, E.O. Aboh & N. Smith (eds), 345–363. Amsterdam: John Benjamins. doi:10.1075/cll.35.21ans

Bailey, R.W. 1998. Majority language, minority misery. In *Language Legislation and Linguistic Rights*, D.A. Kibbee (ed.), 206–224. Amsterdam: John Benjamins. doi:10.1075/impact.2.14bai

Bishop, E. 1983. *The Complete Poems* 1926–1979. New York NY: Farrar, Straus & Giroux.

Bloom, D. 1986. The English language in Singapore: A critical survey. In *Singapore Studies*, B.K. Kapur (ed.), 337–458. Singapore: Singapore University Press.

Braddell, T. 1853. Notes of a trip to the interior from Malacca. *Journal of the Indian Archipelago and Eastern Asia* 7: 73–104.

Burah, T.A. 2006. *Saga of the Exiled Royal Javanese Unearthed*. Dehiwala, Sri Lanka.

Documentation of Endangered Languages (DOBES). 2000–2015. *Sri Lanka Malay*. DOBES/ Volkswagen Stiftung. <http://dobes.mpi.nl/projects/slm/>

Eades, D. 1988. They don't speak in an Aboriginal language, or do they? In *Being Black: Aboriginal Cultures in 'Settled' Australia*, I. Keen (ed.), 97–115. Canberra: Aboriginal Studies Press.

Earl, G.W. 1837. *The Eastern Seas or Voyages and Adventures in the Indian Archipelago in 1832– 33–34*. London: Allen & Co.

Endangered Language Alliance. 2012. Endangered Language Alliance. <http://elalliance.org/>

Fernando, C. 1996. The post-imperial status of English in Sri Lanka 1940–1990: From first to second language. In *Post-Imperial English*, J.A. Fishman, A.W. Conrad & A. Rubal-Lopez (eds), 485–511. Berlin: Mouton de Gruyter.

Fishman, J.A. 1991. *Reversing Language Shift: Theoretical and Empirical Foundations of Assistance to Threatened Languages*. Clevedon: Multilingual Matters.

Garrett, P.B. 2006. Contact languages as "endangered" languages: What is there to lose? *Journal of Pidgin and Creole Languages* 21(1): 175–190. doi:10.1075/jpcl.21.1.05gar

Giles, H., Bourhis, R.Y. & Taylor, D.M. 1977. Towards a theory of language in ethnic group relations. In *Language, Ethnicity and Intergroup Relations*, H. Giles (ed.), 307–348. London: Academic Press.

Gunasekera, M. 2005. *The Postcolonial Identity of Sri Lankan English*. Colombo: Katha.

Gwee, T.H.W. 1993. *Mas Sepuloh: Baba Conversational Gems*. Singapore: Armour.

Gwee, T.H.W. 2006. *A Baba Malay Dictionary*. Singapore: The Peranakan Association; Tuttle.

Haspelmath, M. 2001. The European linguistic area: Standard Average European. In *Language Typology and Language Universals: An International Handbook*, Vol. 2. [Handbücher zur Sprach- und Kommunikationswissenschaft 20(2)], M. Haspelmath, E. König, W. Oesterreicher & W. Raible (eds), 1492–1510. Berlin: Walter de Gruyter. doi:10.1515/9783110171549.2

Heller, M. 2013. Repositioning the multilingual periphery: Class, language and transnational markets in Francophone Canada. In *Multilingualism and the Periphery*, S. Pietikainen & H. Kelly-Holmes (eds), 17–34. Oxford: OUP. doi:10.1093/acprof:oso/9780199945177.003.0002

Hinton, L. 2011. Revitalisation. In *The Cambridge Handbook of Endangered Languages*, P.K. Austin & J. Sallabank (eds). Cambridge: CUP.

Ho, M. L. & Platt, J. 1993. *Dynamics of a Contact Continuum: Singaporean English*. Oxford: Clarendon Press.

Hornsby, M. & Vigers, D. (eds). 2013. *Breton: The Postvernacular Challenge. Special issue, International Journal of the Sociology of Language* 223.

Hussainmiya, B.A. 1986. Melayu Bahasa: Some preliminary observations on the Malay creole of Sri Lanka. *Sari* 4(1): 19–30.

Hussainmiya, B.A. 1987. *Lost Cousins: The Malays of Sri Lanka*. Kuala Lumpur: Universiti Kebangsan Malaysia.

Hussainmiya, B.A. 1990. *Orang Rejimen: The Malays of the Ceylon Rifle Regiment*. Kuala Lumpur: Universiti Kebangsan Malaysia.

Kwachka, P. 1992. Discourse structures, cultural stability, and language shift. *International Journal of Society and Language* 93: 67–73.

Kwan-Terry, A. 2000. Language shift, mother tongue, and identity in Singapore. *International Journal of the Sociology of Language* 143: 85–106. doi:10.1515/ijsl.2000.143.85

Kwok, K.W. 2000. Singapore. In *The Encyclopedia of the Chinese Overseas*, L. Pan (ed.), 200–217. Singapore: Chinese Heritage Centre; Cambridge MA: Harvard University Press.

Ladefoged, P. 1992. Another view of endangered languages. *Language* 68(4): 809–811. doi:10.1353/lan.1992.0013

Lim, B. K. 1917. The Chinese in Malaya. In *Present Days' Impressions of the Far East and Prominent and Progressive Chinese at Home and Abroad: The History, People, Commerce, Industries and Resources of China, Hong Kong, Indo-China, Malaya and Netherlands India*, W. Feldwisk (ed.), 875–882. London: Globe Encyclopedia.

Lim, L. 2010a. Peranakan English in Singapore. In *The Lesser-Known Varieties of English*, D. Schreier, P. Trudgill, E.W. Schneider & J.P. Williams (eds), 327–347. Cambridge: CUP. doi:10.1017/CBO9780511676529.018

Lim, L. 2010b. Migrants and 'mother tongues': Extralinguistic forces in the ecology of English in Singapore. In *L. English in Singapore: Modernity and Management*, Lim, A. Pakir & L. Wee (eds), 19–54. Hong Kong: Hong Kong University Press. doi:10.5790/hongkong/9789888028436.003.0002

Lim, L. 2011. Tone in Singlish: Substrate features from Sinitic and Malay. In *Creoles, Their Substrates and Language Typology* [Typological Studies in Language 95], C. Lefebvre (ed.), 271–287. Amsterdam: John Benjamins. doi:10.1075/tsl.95.16lim

Lim, L. 2013. The politics of English (and Sinhala and Tamil) in Sri Lanka: *Kaduva* of privileged power, tool of rural empowerment? In *The Politics of English in Asia: Language Policy and Cultural Expression in South and Southeast Asia and the Asia Pacific*, L. Wee, R. Goh & L. Lim (eds), 61–80. Amsterdam: John Benjamins. doi:10.1075/wlp.4.06lim

Lim, L. 2013–2016. *LinguisticMinorities.HK*. <http://linguisticminorities.hk>

Lim, L. 2014. Yesterday's founder population, today's Englishes: The role of the Peranakans in the (continuing) evolution of Singapore English. In *The Evolution of Englishes* [Varieties of English around the World G49], S. Buschfeld, T. Hoffmann, M. Huber & A. Kautzsch (eds), 401–419. Amsterdam: John Benjamins. doi:10.1075/veaw.g49.23lim

Lim, L. 2016. Multilingual mediators: The role of the Peranakans in the contact dynamics of Singapore. In *Multilingualism in the Chinese Diaspora World-Wide*, L. Wei (ed.), 216–233. Abingdon: Routledge.

Lim, L. & Ansaldo, U. 2006. Keeping Kirinda vital: The endangerment-empowerment dilemma in the documentation of Sri Lanka Malay. *Amsterdam Centre for Language & Communication Working Papers* 1: 51–66.

Lim, L. & Ansaldo, U. 2007. Identity alignment in the multilingual space: The Malays of Sri Lanka. In *Linguistic Identity in Postcolonial Multilingual Spaces*, E.A. Anchimbe (ed.), 218–243. Newcastle upon Tyne: Cambridge Scholars.

Lim, S. 1988. Baba Malay: The language of the 'Straits-born' Chinese. In *Papers in Western Austronesian Linguistics No. 3*, H. Steinhauer (ed.). Canberra: Department of Linguistics, Research School of Pacific Studies, Australian National University.

Liu, G. 1999. *Singapore: A Pictorial History 1819–2000*. Singapore: National Heritage Board and Editions Didier Miller.

Mufwene, S.S. 2001. *The Ecology of Language Evolution*. Cambridge: CUP. doi:10.1017/CBO9780511612862

Mufwene, S.S. 2004. Language birth and death. *Annual Review of Anthopology* 33: 201–222. doi:10.1146/annurev.anthro.33.070203.143852

Mufwene, S.S. 2007. Population movements and contacts: Competition, selection, and language evolution. *Journal of Language Contact* 1: 63–91. doi:10.1163/000000007792548332

Mufwene, S.S. 2008. *Language Evolution: Contact, Competition and Change*. London: Continuum.

Nathan, J. 1922. *The Census of British Malaysia*, 1921. London: Waterloo & Sons.

Pakir, A. 1986. A Linguistic Investigation of Baba Malay. PhD dissertation, University of Hawai'i.

Peranakan Association Singapore. 1994–2008. *The Peranakan*. (newsletter). Singapore: The Peranakan Association.

Reershemius, G. 2009. Post-vernacular language use in a Low German linguistic community. *Journal of Germanic Linguistics* 21(2): 131–147. doi:10.1017/S1470542709000221

Reid, A. 2000. *Charting the Shape of Early Modern Southeast Asia*. Chiang Mai: Silkworm Books.

Ricklefs, M.C. 1974. *Jogjakarta under Sultan Mangkubumi, 1749–179. A History of the Division of Java*. London: OUP.

Romaine, S. 2014. Linguistic diversity, sustainable development and the economics of language policy. Landau: Talk delivered at *36th International LAUD Symposium*, 31 March–3 April.

Rudolph, J. 1998. *Reconstructing Identities. A Social History of the Babas in Singapore*. Aldershot: Ashgate.

Saldin, B.D.K. 2000. *A Guide to Malay*. Kurunegala, Sri Lanka.

Saldin, B.D.K. 2001. *The Sri Lankan Malays and their language*. Kurunegala, Sri Lanka.

Saldin, B.D.K. 2003. *Portrait of a Sri Lankan Malay*. Kurunegala, Sri Lanka.

Saldin, B.D.K. & Lim, L. 2007. *Kamus Bahasa Melayu Sri Lanka*. A Sri Lanka Malay–Malay–English Dictionary. Frankfurt: Volkswagen Stiftung/DoBeS.

Shandler, J. 2006. *Adventures in Yiddishland: Postvernacular Language and Culture*. Berkeley, CA: University of California Press.

Siemund, P., Gogolin, I., Schulz, M.E. & Davydova, J. (eds). 2013. *Multilingualism and Language Diversity in Urban Areas* [Hamburg Studies on Linguistic Diversity 1]. Amsterdam: John Benjamins. doi:10.1075/hsld.1

Song, O.S. 1923. *One Hundred Years' History of the Chinese in Singapore*. London: John Murray. Reprinted 1967, Singapore: University of Malaya Press.

Sourjah, M.A. 2003. Malays in Sri Lanka: Came with their women-folk. *WWW Virtual Library – Sri Lanka*. 8 January 2004. <http://www.lankalibrary.com/cul/malays.htm>

Tan, C.B. 1988a. *The Baba of Malacca. Culture and Identity of a Peranakan Community in Malaysia*. Petaling Jaya, Selangor: Pelanduk Publications.

Tan, C.B. 1988b. Structure and change: Cultural identity of the Baba of Melaka. *Bijdragen tot de Taal, Landen Volkenkunde* 144: 297–314. doi:10.1163/22134379-90003298

Thaliph, M.F. 2003. *Book on Malay Grammar (Nahu): A Guide to Write and Speak Grammatical Malay*. Wattala, Sri Lanka.

Thomson, J.T. 1875. *The Straits of Malacca, Indo-China and China, or Ten Years' Travels, Adventures and Residence Abroad*. London: Sampson Low; Marston: Low & Searle.

Vaughan, J.D. 1879. *Manners and Customs of the Chinese in the Straits Settlements*. Singapore: The Mission Press.

Whorf, B.L. 1941. The relation of habitual thought and behavior to language. In *Language, Culture, and Personality: Essays in Memory of Edward Sapir*, L. Spier, A.I. Hallowell & S.S. Newman (eds), 75–93. Menasha: Sapir Memorial Publication Fund, Reprinted in 1956, in Carroll, J. B. (ed.), *Language, Thought and Reality: Selected Writings of Benjamin Lee Whorf,* 134–159. Cambridge MA: The MIT Press.

Woodbury, A.C. 2005. Ancestral languages and (imagined) creolisation. *Language Documentation and Description* 3: 252–262.

Yap, K.H. 2008. Here's re-looking at you, Bibik. *The Straits Times*, 2 February 2008.

Reacting to language endangerment

The Akie of north-central Tanzania

Bernd Heine, Christa König and Karsten Legère
University of Cologne / University of Frankfurt / University of Gothenburg

Language endangerment in Africa is of a different kind than it is in many other parts of the world. Globalization and the impact of languages such as English, French, or Portuguese are not a major problem for the maintenance of African languages. Language loss is no less a factor in Africa than it is elsewhere in the world, but the replacing languages are, with very few exceptions, not international languages such as those of the former colonial powers but rather fellow African languages (Sommer 1992; Brenzinger 2007a: 197; 2007b).[1]

On the basis of the research findings on language endangerment that exist it is possible to understand the main factors that can be held responsible for language replacement. But there is less information on why replacement does *not* take place – that is, why people in certain situations do not give up their heritage language in favour of some other language even if their sociolinguistic environment discourages such a behavior. This is the question looked into in the present paper, using the Akie, a traditional hunter-gatherer people in Tanzania as an example.

Keywords: African, Akie, defensive culture, language purism, language transmission, Maasai, quadrilingualism, Swahili

1. Introduction

1.1 Language death in Africa

Language endangerment in Africa is of a different nature than in many other parts of the world. Globalization and the impact of languages such as English or French

---

1. While this issue is in need of much further research, it would seem that one reason can be seen in the fact that languages such as English or French tend to be acquired in formal education rather than in informal, everyday communication in most parts of Africa.

Table 1. Status of West African languages (adapted from Blench 2007: 143)

Status	Number of languages	Percentage
Not threatened	683	63.5
No information	304	28.3
Definitely threatened	55	5.1
Probably extinct	16	1.5
Declining	10	0.9
Moribund	7	0.7
Total	1075	100.0

is not a major problem for the maintenance of African languages. Language loss is not less a factor in Africa than elsewhere, but the replacing languages are as a rule not international languages such as those of the former colonial powers but rather fellow African languages (Sommer 1992; Brenzinger 2007a: 197).

Considering that half of the world's languages are said to have disappeared in the course of the last 500 years, Africa's rate of language loss is remarkable, but not really dramatic.

On the basis of a survey of 1075 languages of West Africa, Blench (2007) proposed a classification of the languages based on their relative degree of endangerment. As the data provided by Blench (2007) suggest, nearly two thirds of the languages of West Africa are not threatened and only less than ten percent are clearly endangered (see Table 1). On the basis of such observations one may predict that the majority of the roughly 2000 African languages are not immediately endangered, that is, are likely to survive the present century.

A wider perspective of sociolinguistic situations in Africa suggests, however, that such predictions must be taken with care. On the one hand, there are a number of African languages that have disappeared over the last century. On the other hand, various cases have been reported where earlier writers had predicted that a given language will soon be extinct yet where the language concerned is still alive and well to this day (cf. the data in Brenzinger 1992, 2003; Sommer 1992; see also Petrollino & Mous 2010).

Why then do languages in Africa die out? A number of reasons have been proposed. Perhaps the factor most frequently invoked is the number of speakers and the minority status of a language. Another frequently named factor concerns the prestige of a language, even it is not always entirely clear what exactly this notion stands for. Batibo (1992) argues that it is the relative degree of language prestige that "has been the major determinant of language shift in Tanzania," and he proposes the following hierarchy (see also Legère 1992):

(1) A scale of types of prestige hypothesized to determine language shift in Tanzania
  1   Language of national prestige
  2   Language of regional prestige
  3   Language of local prestige
  4   Language without special prestige[2]

Rather than such language-internal factors, external factors have also been proposed. A cursory survey of language shift situations in Africa suggests that quite commonly it was one or more of the following factors that appear to have played a role: (a) the impact of a dominant culture, (b) a global ideology or religion, and (c) the transition from one form of economy to another.

But factors such as these are neither necessary nor sufficient for language shift. There are some linguistic communities in Africa that have resisted language replacement. What such observations suggest is that language development is hard to predict in Africa. The question to be looked into in this paper is the following: Why does a community that has been predicted to give up their language nevertheless not do so?

The reasons are complex, we are restricted here to one kind of reason – one that can be described with reference to the notion "defensive culture".[3] In the remainder of the paper we will illustrate the significance of this notion by looking in more detail into the situation of what has been portrayed in previous research as a seriously endangered language. This is the Akie language of Tanzania.

1.2   A "defensive culture"

In the 1920s, the British government official R. Maguire, placed in Kibaya (Kiteto District), studied the Akie hunter-gatherer community of north-central Tanzania.

---

2. However, as long as ethnic and linguistic homogeneity prevails, the question of language prestige seems to be rather insignificant in a number of cases. Only when language contacts and ethnic-linguistic diversity are relevant for an area, the question of which language could facilitate the communication among linguistically different people arises. Often in what is now Tanzania the answer was Swahili for its wide-spread distribution as a second language and, in more recent years, its status as (semi-) official language used in administration and in other domains.

3. An anonymous reader of this paper suggests the terms 'exclusive', 'protective', and 'sheltering' culture, instead. To the extent that the people that are the main concern of this paper aim at defending institutions that are conceived as central to their socio-cultural life against influence from outside we prefer the notion "defensive culture" for the purposes of the following discussion.

He concluded that these people (called Mósiro by him) were seriously endangered in their linguistic identity:

> All the Mósiro speak Masai, but many of them do so very imperfectly … The language of the *Mósiro* is dying … I have asked many *Mósiro* to give me the names of various common objects, and I have often been given a Masai name, my informant protesting that he knew no other.   (Maguire 1948: 10)

On the basis of such an assessment one might expect that by now, nearly a century later, this language is no longer alive. More than half a century after Maguire, researchers such as Rottland (1982), Kaare (1996) and Bakken (2004) found Akie still to be spoken, even if they classified it as seriously or critically endangered (Brenzinger 2007a: 199) or as "probably dying out", to be replaced by Maasai (Sommer 1992: 305; Legère 2012; MLT 2009; Heine, König & Legère 2014; see also Rottland & Voßen 1977).[4]

There would in fact be reasons for the language to have disappeared: First, Akie is spoken presumably by hardly more than 300 people whereas the language of the surrounding Maasai has well over one million speakers. Thus, in terms of their absolute number of speakers, the Akie are a tiny minority. Second, in the ideology of the Maasai, acknowledged to some extent by Akie people, the latter are socially, culturally, and economically a fairly marginal group.[5]

Third, most Akie have a fluent command of Maasai, and for many of them it is the primary language (see Table 2 below). Fourth, the Akie are economically dependent on the Maasai. Herding the lifestock of the latter and selling them their honey is one of the major income sources for the Akie. Fifth, they are also culturally dependent. For example, they have largely adopted the age-set system of the Maasai and may practice the ceremonies associated with the system jointly with the Maasai. Younger Akie also imitate the dressing and ornamentation practices

---

4. Following Heine (1980), the Maasai are sometimes referred to as Maa since they are part of the Maa society, which in addition to the Maasai also includes other ethnic groups such as Parakuyu (see 1.3.2.1), Arusha, Samburu and Chamus. The first two are restricted to northern Tanzania, the last two live in north-central Kenya, while the Maasai are found in the same way in Tanzania and Kenya.

5. What Kaare (1996: 149) claims with reference to "Dorobo" groups in Kenya and Tanzania (see also Klumpp & Kratz 1993) applies to quite some extent to the Akie, being part of these groups: "… what underlies the Maasai idea about symbiotic relations with their Dorobo neighbours is the construction of marginality". Marginalization affects the Akie at all levels of social and economic life. Akie women find it hard to sell their products on the local markets: "Some of my female informants told me for example about their problems of going to the market and trading commodities. They stated that other people were avoiding their wares because they were assumed … to poison them" (Schöperle 2011: 45).

of the Maasai. And finally, for a number of Akie, living like a Maasai pastoralist would be clearly preferable to leading the fairly miserable existence of an Akie hunter-gatherer.

Thus, there would be reasons enough for language shift to take place. As a matter of fact, however, this does not clearly appear to happen: As we will see in more detail later, Akie is still spoken today. [6] We will argue that the main reason of the Akie people for not having given up their language is that they have developed a "defensive culture".

With this term we are referring to speech communities that make a deliberate attempt to keep their language and culture separate from those of surrounding cultures. This attempt manifests itself in particular in the following features of behavior:

(2) Manifestations of defensive behavior
   a. *Traditional culture:* Members of other communities are strongly discouraged to participate in inherited cultural practices, such as rituals, dances, singing, etc.
   b. *Language use:* The language is used in some sense as a "secret code".
   c. *Language planning:* There is a pronounced "purist" attitude whereby attempts are made to keep the language "clean", i.e., free from lexical and other material from other languages, even if the Akie are not really successful in these attempts, as the multitude of Maasai borrowings and code-switches suggests.

1.3   The present paper

Based on a field survey carried out by the present authors, this paper looks into the question of what the present state of a language is that ninety years ago was believed to be in a process of extinction.[7] The paper is based on field surveys carried

---

6.   We are concerned here with the Akie of Kilindi District, that is, the eastern part of the Akie territory. That the situation is different in the western part (Kiteto District) is suggested by observations made by Schöperle (2011), who found hardly any Akie speakers in this district: "… the Akiek-language is only spoken by a small minority of elderly people. Some of my informants had the ability to understand it but I found only two adult men who were able to communicate fluently in it. The younger generations meanwhile have adopted Maa as their mother tongue, and Maa is in fact the most widely distributed language throughout the region" (Schöperle 2011: 35–36).

7.   In this survey, demographic and sociolinguistic information was collected on 115 people who claimed, or were claimed by others, to belong to the Akie community. Whether, or to what extent, this sample is representative of the Akie as a whole is open to question; we suspect that

out within the *DoBeS* (Documentation of Endangered Languages) program of the *VolkswagenStiftung* (Volkswagen Foundation) between January and March, 2013 and 2014.[8] The primary concern in this project is with language documentation, and this paper is restricted to findings made on the sociolinguistic situation of the Akie language.

2.   Akie and language endangerment

The Akie-speaking people inhabit mainly the Kilindi District of Tanga Region and the Kiteto and Simanjiro Districts of Manyara Region of north-central Tanzania between Handeni to the east and Kibaya and Simanjiro to the west; they seem to have been living in this area before all their neighbors, such as the Maasai or Nguu (Ngulu), arrived (Kaare 1996; Bakken 2004: 38ff.; Schöperle 2011). The language is fairly homogeneous, there do not appear to be any dialects or other noticeable linguistic cleavages, even if phonetic variation among speakers is quite pronounced. The language is not used in writing.

The Akie call themselves *akie* (singular *aki-ántee*). According to Rottland (1982: 305; p.c.), their autonym is *ákiék* (with unreleased final *k*). As for the number of Akie speakers, Sommer (1992: 305) mentions a figure of less than 1000 speakers, and Brenzinger (2007: 199) of 50 speakers, both classifying it as a seriously or critically endangered language. According to our own estimate, it is hardly more than 300 people who still speak Akie.

The Akie are known in Tanzania as (*N*)*dorobo*, a term originating from the Maasai noun *ol-tórobóni* (sg.), *il-tórobo* (pl.) 'people without cattle, poor people'. The term, also applied to other traditional hunter-gatherer groups in the area, such as Kisankare (*kisaŋkárɛ*), Kinyalang'ate (*kiɲalaŋáte*), and the Cushitic

---

language Aasá (Áasax, Aramaní, Aramanik), clearly having derogatory connotations but used by the Akie people themselves vis-à-vis other ethnic groups.

Earlier information about Akie falls under the rubric of research on "Ndorobo", "Nderebo", "Dorobo", or Mósiro (Maguire 1948; Maghimbi 2005; Legère 2006). The only readily available linguistic material stems from Rottland (1982). What Rottland's work demonstrates beyond reasonable doubt is that Akie is a language closely related to the cluster of Kalenjin languages spoken in Kenya, Northern Tanzania, and Southeastern Uganda, including the Kenyan traditional hunter-gatherer communities commonly referred to as Okiek. The Akie are, however, essentially unaware of this relationship;[9] there are no contacts of any kind with their linguistic relatives in Kenya and Uganda, or in Tanzania.

Being a member of the Kalenjin cluster, Akie belongs to the Southern Nilotic branch of the Nilotic family (see Rottland 1982). The latter has been classified as belonging to the Eastern Sudanic branch of the Nilo-Saharan phylum (Greenberg (1963).

According to a sociolinguistic survey of 115 Akie people that we were able to carry out between January and March, 2013, the majority of these people were essentially quadrilingual. As Table 3 shows, nearly two thirds of them have not only Akie but also Maasai as their primary languages, and the Bantu languages Nguu (or Ngulu)[10] and Swahili are not only useful secondary languages but are also spoken by a sizable portion of the population as primary languages – that is, as languages spoken with close to mother tongue competence.

After Maguire (1948) and others had observed indications of a process of language decay, the question then is: Is Akie presently threatened by extinction? One way of looking into this question is by using the list of factors proposed by UNESCO (2003) in its document *Language Vitality and Endangerment* for assessing language endangerment. These factors are listed in Table 4 (Brenzinger 2007: 544–555).[11]

We will now look at each of these factors in turn.

---

9. This has changed, however, since we stayed with them. As we learned from recent visitors to them, our informing them about their linguistic ancestry had an impact on their understanding of their past: At least the Akie with whom we worked are now convinced that the Kalenjin people in Kenya, speaking a closely related Southern Nilotic language, are their cousins.

10. Nguu (or Ngulu, G31 according to the reference system of Bantuists) belongs to the Zigula-Zaramo group of Bantu languages (Guthrie 1948; Nurse & Philippson 2003).

11. The UNESCO document adds a ninth factor to the list of Table 4, namely "Urgency of documentation: Amount and quality of documentation". This factor is omitted here because it is of a different nature, relating to what needs to be done rather than what there is.

Table 3. Languages known by the Akie people
(based on König et al. 2014, Appendix 1)[12]

Age group	Their primary language	A secondary language	No knowledge	Total in percent	Total number of persons
Akie	61.1%	11.5%	27.4%	100	113
Maasai	65.2%	14.3%	20.5%	100	112
Nguu	49.0%	22.1%	28.9%	100	104
Swahili	38.4%	49.5%	12.1%	100	99
English	0%	0%	100%	100	113
Other languages	0%	2.4%	97.6%	100	86

Table 4. Degree of language endangerment according to UNESCO
(2003; see also Brenzinger 2007 and Legère 2008 for discussion)

1. Intergenerational language transmission
2. Absolute number of speakers
3. Proportion of speakers within the total population
4. Loss of existing language domains
5. Response to new domains and media
6. Material for language education and literacy
7. Governmental and institutional language attitudes and policies, including official status and use
8. Speakers' attitudes towards their own language

1. *Intergenerational language transmission.* For good reasons, this factor is widely accepted to be the most crucial to assess language endangerment (Brenzinger 2007: 545; Legère 2008). Presumably the most reliable indicator of possible language replacement is found in the language behavior of young people. In our survey we were restricted to a comparison of two age groups, namely people below 20 years of age and people from 20 years onward. Table 5 summarizes the patterns of language knowledge among these two age groups.

While Akie is usually transmitted by parents to their children, it is an endangered language. As observed earlier, Maguire predicted language replacement some 90 years ago, and Brenzinger (2007: 199) classifies Akie as a seriously endangered language, that is, as one that is spoken only by the grandparental generation and above (see Krauss 2007: 1). As can be seen in Table 2, this is not entirely in accordance with the findings made in our survey: Akie is spoken as a primary

---

12. The findings presented here and in the following tables are based primarily on self-assessments. Only in a smaller number of cases could they be tested by means of participant observation.

Table 5. Numbers of Akie having a knowledge of Akie according to age
(based on Appendix 1 of König et al. 2014)

Age group	Their primary language	A secondary language	No knowledge	Total in percent	Total number of persons
20 and over	73.1 %	11.5 %	15.4 %	100	78
2–19	34.3 %	11.4 %	54.3 %	100	35
Total	61.1 %	11.5 %	27.4 %	100	
Total	69	13	31	100	113

language by nearly three quarters of the adult population, even if hardly more than one third of the people below 20 years have a full command of it, and more than half of the young population (54.3% of Table 3) are ignorant of it altogether. Accordingly, it would seem more appropriate to classify Akie as "instable" or "eroded" in the sense of Krauss (2007:1), that is, as one that is still spoken by at least some children.

Does this mean that Akie people are in a process of shifting from their own language to Maasai? The figures in Table 6 suggest that there is no clear answer to this question. If the figures can be taken to reflect a change from one generation of speakers to another then they suggest that, in much the same way as Akie is losing speakers, Maasai is gaining speakers, being one of the primary languages of four out of six young Akie – that is, Maasai appears to be the most important means of communication for younger Akie. For more evidence in support of this hypothesis (see König et al. 2015).

Nevertheless, there is no clear indication that Maasai will replace Akie. First, for about a third of all younger Akie (34.3%, Table 5), their own language is still a primary language. And second, almost one third of younger Akie have no knowledge of Maasai. On the basis of such observations it would seem hard to predict how attitudes on language behavior will develop in the years to come.

To conclude, there clearly is intergenerational language transmission, even if there is also a possible shrinking process: The percentage of younger Akie (34.3%) is less than half that of generations above 20 years of age (73.1%). This fact could be indicative of loss in transgenerational transmission, but need not be: Conceivably, young Akie learn the language of their parents later in age. We observed in fact a few cases where this happened, but these observations were far from sufficient to allow for any predictions.

2. *Absolute number of speakers.* As observed above, there are hardly more than three hundred Akie speakers.[13] Compared to the number of well over one million Maasai people surrounding them, this figure suggests that the Akie form a tiny linguistic minority within an essentially Maasai-speaking world.

3. *Proportion of speakers within the total population.* Table 3 provides an approximate answer to this question: About two thirds of the Akie community speak their language at least as one of their mother tongues.[14] With reference to Tanzania as a whole, the number of Akie speakers is insignificant.

4. *Loss of existing language domains.* A comparison of the Akie (or Mósiro) people of roughly 90 years ago with their modern descendants suggests that none of the existing language domains was really lost: Like their ancestors, the Akie are still predominantly hunter-gatherers (see König et al. 2014), even if their farming activities have become more prominent.[15]

5. *Response to new domains and media.* Obviously, there is response to new domains, but these domains have hardly affected their economy and culture. These new domains concern in particular (a) new techniques of growing maize and beans, (b) consumption, e.g. new kinds of alcoholic drinks, spices, and other goods, and (c) material culture (tools, etc.). Communication with the outside world is severely limited: Cell phones or transistor radios, forming part of the every-day life of modern Tanzanians, are hardly available to the Akie.

Linguistic communication on such innovations may take place in Akie but, more likely, will rely on any of the three other languages they are familiar with,

---

13. This figure is fairly vague, for the following reason: Quite a number of Akie live far away from the settled areas of Kilindi and Kiteto Districts deep in the forest and savannah areas of the Maasai Steppe of north-central Tanzania, inaccessible even to other Akie (as well as the anti-poaching units of the national authorities). We were able to get second-hand information at least on some of them from fellow Akie but were assured by them that there were more Akie people in the forests than they could tell.

14. We are ignoring here the issue of alternative definitions. For our Akie consultants in particular, Akie include in a wide sense all those that subscribe to a traditional hunter-gatherer culture, even if they may now be primarily farmers of pastoralists, and their mother tongue may be Maasai, Nguu or any other language. Thus, Kisankare, Kinyalang'ate, or Aramani people are also clear "Akie" (see König et al. 2014). With our use of the term "Akie" we are restricted to Akie speakers and their immediate relatives.

15. The situation is different in larger parts of Kiteto District, where has been a large-scale shift from a forager to a food-producing economy, accompanied by language shift (see Schöperle 2011 for details).

namely Swahili, Maasai, Nguu and, most probably, Swahili, which has the most elaborated lexical means for expressing relevant concepts.

6. *Material for language education and literacy.* Akie is an exclusively oral language. There are no kinds of written documents that could serve as a basis of education and literacy and, what is more, there is no orthography that could be used.[16]

7. *Governmental and institutional language attitudes and policies, including official status and use.* The Akie and their language are distinctly beyond the scope of national or regional political or social concern. At least for some people in Northern Tanzania there are people known as "Ndorobo", believed to adhere to some form of archaic subsistence. But there is hardly any information on who the "Ndorobo" are, whether they have a language of their own -- or whether they exist in the first place.

8. *Speakers' attitudes towards their own language.* Speakers' attitudes towards their own language are complex. On the one hand, they realize that their language has serious deficits:

a.   Its communicative value is low, being restricted to the small number of its speakers.
b.   To express new cultural, economic, or technological concepts it has to rely on borrowings from Maasai and Swahili.
c.   Speakers realize that Akie with its tiny speech community is hard, if not impossible, to find a generally recognized role in the modern nation state of Tanzania.

At the same time, speakers adhere to the language, for the following reasons:

a.   Being still largely a hunter-gatherer community, there would be no other language that could be used to express the wide range of lexical and constructional means that is needed to survive in this cultural and economic niche.
b.   Akie society includes in much the same way the living and the dead, that is, the *asíiswe* 'ancestors'; Schöperle (2011:97) was told by his Akie informants in Kiteto district "that without remembering the life of the ancestors and the highlighting [of] the hunt the spirit of the Akie would die out". No major decision can be made without the approval of the *asíiswe* (see König et al. 2014). Addressing them in any language other than Akie is hard to conceive for the

---

16. In 2013 we prepared a small primer for Akie willing to read and write in their own language. Whether this first piece of writing will have any impact on them remains to be seen.

people and, what is more, would be strongly disapproved by the *asiiswe*, resulting in misfortune of all kinds.[17]

c.   All attempts that Akie people may have made in the past to become part of Maasai society, or at least to be accepted as social equals have failed. Maasai commonly take Akie wives, but these women will never become Maasai, they will be stigmatized for their whole life as Akie, that is, as second-class members of their new family. The Akie have reacted to all this by setting themselves off from their neighbors and strengthening their own socio-cultural identity (Kaare 1996), and their language is for them a paradigm symbol of this identity; we will return to this issue in Section 4.

Conclusions. According to the description volunteered by Maguire (1948), Akie was clearly more endangered than it is now. While he does not provide detailed information, his description suggests that about 90 years ago the language was on the verge of being replaced by Maasai. This situation appears to have changed: First, the language continues to be transmitted to children. Second, it has not lost any of the domains of use that are central to their present life and, third, the majority of Akie people want their language to be retained.

But what accounts for this apparent reorientation in their attitudes that the Akie appear to have experienced in the course of the last 90 years? We will look for an answer in the next section.

3.   Akie as a "defensive culture"

In (2) of Section 1.2 we proposed three main kinds of manifestation for defining a "defensive culture", and it would seem that Akie culture can in fact be viewed as an example of it. We will now look at the factors with reference to their significance for the Akie.

3.1   Traditional culture

The Akie view Maasai people as their close cousins with whom they share the same origin, and consider them to be part of their cosmos; one of the explanations given for why they dress in a way similar to the Maasai is because they "all came from the same honey bag" (Kaare 1996: 179). But the attitudes that the Akie

---

17.   Such beliefs and activities appear to apply throughout the male adult population of the Gitu area, where we carried out our field research. Whether, or to what extent, they apply to the Akie community as a whole could not be tested during our field research.

have vis-a-vis their numerically, economically, and socially much more powerful neighbors are complex. There is on the one hand admiration but, on the other hand, there is also suspicion and mistrust. We may illustrate the latter with a couple of examples.

There were a number of social functions that we witnessed where Akie people tried to keep Maasai people away from participating in or getting information on inherited cultural practices such as rituals, dances, singing, etc.

One aspect of these relations concerns what the Akie portray as protection of their cultural life. This behavior surfaced time and again during our work; a few observations may illustrate it. When a group of about a dozen Akie people were celebrating the Blessing-the-Hunting-Weapons ceremony in February, 2013, there was a Maasai woman watching. The Akie would then interrupt their activities and insist that that woman leave before they would continue.

In the same month there was another Akie dancing festivity also taking place at the fringe of the forest close to the village of Gitu. When two Maasai men approached, the dancers stopped and there was a small uproar, characterized by passionate discussions. In the end, the Akie told us that the Maasai would destroy their ceremony. Therefore, dancing would only be resumed when the Maasai had left, and this is what then happened.

In another event in January, 2014, a group of nine Akie gave us a demonstration of how they would ritually demarcate a plot that could not be approached by enemies or dangerous animals (the *mundee* ceremony). For some time there was a middle-aged Maasai man present, known to the Akie people who were singing and dancing. At some point one of the Akie men jumped up and physically threatened the Maasai, blaming him for trying to "steal" their culture. It was only when the Maasai man was removed from the place that singing and dancing resumed.

Such observations, impressionistic as they are, were a common theme during our research among the Akie. The people would freely share most central aspects of their ritual practices with the alien researchers but not with the Maasai, in spite of the fact that the two peoples share a long history of symbiotic interaction (Kenny 1981; Kaare 1996), and at times even friendship (Maguire 1948).[18]

Note that they never had problems with non-Maasai people being present, which frequently happened during our work, nor with us. Being aware that our

---

18. Symbiosis manifested itself most clearly but not only in situations of famine. When the Maasai were close to being wiped out by the rinderpest epidemic in the 1890s, the Akie helped them to survive by providing them with game meat and, conversely, the Maasai supplied the Akie with livestock products when the latter were unable to hunt (Maguire 1948). Furthermore, at times in the past when the ivory trade was thriving the Akie killed many elephants for Maasai to sell the tusks to outside traders.

goal was to document their language and culture they were at all times highly cooperative, even revealing details of cultural activities that were inaccessible to their own women and children.

3.2   Language use

The language is used in some sense as a "secret code", Akie people try to keep the Maasai away from it. Working on the western Akie in Kiteto District, Marianne Bakken (2004:150) noted:

> […] they (the Akie) usually even refrain from speaking the Akie language at all when within hearing distance of their neighbors, whether Bantu or Maasai. This became apparent to me on our visits to the monthly market in Kibaya town, and in the larger shopping villages. When friends and kin meet and want to exchange news, they typically separate themselves spatially from non-Akie by clustering in small groups on the outskirt of the market place, or under a tree where they talk together in a low-voiced manner. In fact they only speak Akie freely in the secrecy and privacy of their own homesteads and settlements, or when alone in the bush.

The following example may illustrate a common attitude shown by them. When doing our field work in front of our tent in Gitu Juu next to an Akie home, people quite commonly used to pass by watching us. One day in January, 2014, there was a Maasai herding boy curiously watching what we were doing. Our Akie consultant Nkoiseyyo became irritated and shouted at the boy, asking him to leave us alone. Since even young Maasai have little respect for the Akie, the boy ignored Nkoiseyyo's words and stayed on. The latter then turned to us: "Why don't you chase this Maasai away, you never know what these people really are after."

We observed similar attitudes time and again. Note that roughly four fifths of the Akie speak Maasai very well (see Table 3). Many of their conversations are suggestive of a high degree of code-switching and, when dealing with certain topics relating to modern life, Akie people, both men and women, may switch entirely from their own language to Maasai, sometimes not even being aware that they are not using their own language.

But this can change rapidly in the presence of Maasai people. For example, when a Maasai person approaches a group of Akie, the latter might stop talking or turn to their own language. And in such a case they tend to make a deliberate effort to talk in a way that is unintelligible to Maasai people, avoiding code-switching and Maasai loanwords, using Akie as a kind of secret code. In such a situation they may turn into "language purists".

3.3   Language "purism"

We will loosely say that language "purism" is an attitude where speakers make an attempt to keep their language free from lexical and other material of another language in spite of the fact that that material is an integral part of their speech.

That there is a pronounced "purist" attitude was apparent when we analyzed the Akie texts we had recorded with our consultants. These texts contain a number of expressions that were beyond any reasonable doubt taken from Maasai.[19] When our Akie consultants came across such expressions they frequently asked us to eliminate them and replace them with corresponding Akie terms. Our submitting that these texts must remain as recorded, hence, could not be altered did not convince them. They would ask why we wanted "mixed" texts rather than "pure" Akie. And even arguing that they themselves had produced these Maasai expressions in the texts had little effect on them.[20]

Such problems never concerned borrowed material from other languages, as if it was only Maasai that was "contaminating" their language. Maasai language and culture was viewed as an eroding force that was on the way to destroying everything that the Akie aimed at defending: A distinct culture, society, and language.

While "purism" is an important ideological factor for many Akie consulted by us, as a matter of fact it is essentially ignored in informal language use, which abounds with lexical material borrowed, mostly from Maasai but also from Swahili.

3.4   Why is there a "defensive culture"?

If the hypothesis is correct that the Akie represent a defensive culture then one may wonder what motivation there may be for such a behavior.

The evidence available suggests the following answer. The Akie realize that there is no place for them in the modern nation state of Tanzania. Except for a few men who are recruited as game scouts or trackers to the national parks of the

---

19. There are also many Swahili loanwords but most of them have entered the Akie language via Maasai.

20. The basis of this disparate behavior between linguistic ideology, on the one hand, and actual language use, on the other, is in need of much further research.

country, especially Tarangire, Selous, and Serengeti, they do not participate in any national activities.[21]

– Their inherited economy is no longer sufficient to feed them and, what is more, is discouraged by national and regional authorities. The natural resources on which this economy is based, and which are needed for physical survival, have dwindled.
– While many nowadays do some farm work, they are clearly less successful in this work than any of the peoples around them.
– Being too few to be recognized as a distinct segment of the national population, they have no political standing to voice their concerns.[22] Note also that they are generally held in low social esteem by all their neighbors.
– Their children do not take advantage of the educational facilities provided by the state.[23]

But more important than all this is the fact that the Maasai have set a social barrier that the Akie cannot cross, even if quite a few would like to, such as:

– Some Akie have in fact acquired livestock and adopted Maasai as their culture and language. Yet, they will never be accepted by the Maasai as being the same.

---

21. Working on the Akie in Kiteto District in the 1996/7, Marianne Bakken writes: "They [the Akie] have limited access to a seat on the village council (total seats 25) as Councillors on the basis of pure numbers. According to my experience their village neighbours also have no interest in voting in 'poor Dorobo' who in their eyes do not share important values with themselves" (Bakken 2004: 169).

22. Already two decades ago, Kaare (1996: 188) observed: "Because their Bantu neighbours control the entire government structure, Akie feel left out of most of the decisions which pertain to village administration. In most cases they are just drafted to work in village self-help schemes from most of which they do not benefit as they spend most of their time away from the village in the forest either hunting or collecting honey or employed as casual labourers."

23. The village of Ngababa (Ngapapa) in the east of Kiteto District is a notable exception: Schöperle (2011: 82) found most of the children in the village school to be Akie from all over the district.

–   Akie girls are commonly married to Maasai men but they never acquire a sta-
    tus that is equal to that of a Maasai woman.[24] Note that, according to Bakken
    (2004: 160), Akie women would prefer Akie to Maasai men.[25]

In this world of socio-political, economic, and cultural isolation, the Akie appear
to have turned to their own traditional world in which they feel at home. In doing
so, they themselves are setting boundaries vis-a-vis their human environment,
trying to keep their culture "clean" and to prevent outsiders, and essentially al-
ways the Maasai, from having access to it.

One effect of this behavior is resistance to language shift. Whether the Akie
will be successful in these efforts, and what their long-term consequences will
be is hard to predict. And it is equally unclear whether, or to what extent, the
situation of the Akie, as portrayed in this paper, can be related to or can offer an
explanatory basis for that of some other linguistic minorities in Africa.

4.   Language, culture and the future

One possible outcome of the research discussed in this paper is that the use of the
Akie language will possibly be extended to some form of written communication.
Another outcome will presumably be that the people's awareness of their language
as a symbol of social and cultural identity will be strengthened. But this raises a
number of problems.

One problem concerns the potential contribution that researchers can make
for the future development of the language and its social functions: It would seem
that such a contribution must remain a modest one. But another, even more se-
rious problem is the following. The Akie language is deeply entrenched in the
traditional culture of the community. The following example may illustrate this.
As we observed in (viii) of Section 2, the life of all Akie persons we came across
is to a large extent determined by their relationship with the ancestors (*asííswe*,
sg. *asííswantee*), and the *asííswe* have to be addressed only in the Akie language,
even if in practice this is not always the case. Once one's father or mother dies s/he

---

24. For Maasai in Kiteto District "marrying an Akie woman recently became very fashionable,
because the bride prices for these girls are generally lower than for Maasai-girls. They told me
that especially young men who possess only few livestock and little money would approach
Akie women. This statement was affirmed by some elderly hunters from Napilo Konya, who
told me that most bride prices for Akie girls would be comparably little, because many families
were economically weak (Schöperle 2011: 43).

25. "Akie women … would rather marry Akie, feeling unsecure as to settle among the Maasai
whatever other benefits it could bring in terms of material wealth" (Bakken 2004: 160).

turns into an *asíiswantee*, meaning that s/he is always watching you – as far away as you may be – one cannot escape them. For example, going to town for a drink in a bar does not help: The *asíiswe* are already there, waiting for their share. The *asíiswe*'s main complaint is that they are hungry. In order not to be hit by accidents or death, people bless the ancestors at all special occasions, especially when something unexpected has happened or one intends to travel. Another complaint of the *asíiswe* is that the Akie are rapidly discarding their traditional culture in favor of a *puuní* (Maasai) or *mɛɛ́yɛ* ("Swahili") life style (König et al. 2014).

One central concern of the *asíiswe* is that the Akie do not give up their economy, and in particular hunting: Most interactions with the *asíiswe* involve hunting in one form or another – be that in personal exchanges or in communal ceremonies, such as the Blessing-the-Hunting-Weapons ceremony. Abandoning this economy would be strongly disapproved by the *asíiswe*, resulting e.g. in misfortune, sickness, and/or death. To conclude, language, ancestor cult and economy are strongly intertwined domains and abandoning one would have repercussions for the other two.

What this means is that any language planning activities aimed at changing the status of Akie are likely to also affect the other two domains of culture. In particular, strengthening the position of the language by developing it into a medium of written communication might entail strengthening the significance of the other two domains as well.

But there does not appear to be much of a future for their traditional economy. Hunting is strongly discouraged by the government[26] and for collecting wild honey, a government licence is required, which is hardly affordable for most Akie people. Accordingly, the people may be facing a situation where they are forced to look for alternatives to earn their living. Attempts in this direction are already on the way: More and more Akie are engaged in farming, maize and beans being their primary crops.

To conclude, the situation of the linguist working on this language is not an easy one: Should s/he focus on documentation work or should s/he aim at strengthening the status of the language and support the aspirations of many people to prevent their language from being lost? Is s/he really sufficiently competent to assess the social, cultural and economic implications that any planning activity towards language empowerment might entail for the future of the Akie society? The responses that we received from the Akie people on this issue do not allow for clearly deciding on this issue. It would seem that there is so far no straightforward answer to these questions.

---

26. There is substantial hunting taking place in the Akie area but rather than by Akie most of it is performed by professional poachers from outside who, unlike the Akie, use firearms.

5.   Conclusions

The patterns of language knowledge and language use among the Akie are complex and it is hard to predict any general trends of development. The most plausible conclusion would be that the Akie people are heading for a situation of quadrilingualism where each of the four languages, namely Akie, Maasai, Nguu, and Swahili, has its place and its preferred functional domain. Akie relates the community to its traditional past, which is far from becoming extinct. Maasai symbolizes the traditional relationship between the Akie and an essentially pastoralist outside world, while Nguu is a symbol of a new economic orientation towards an existence as farmers – an orientation that a number of Akie people see as their main alternative to a gradually declining existence as hunter-gatherers. Swahili finally offers a link to a new world – one that makes it possible for the Akie to participate in activities of a modern African nation state.

While language knowledge thus exhibits roughly a quadrilingual profile, there is one language that enjoys a privileged role, namely Maasai. At the same time, Maasai is also a threatening force and the Akie are aware of this potential and are reacting to it. What the success of their defensive behavior will be remains to be seen.

Acknowledgements

---

27. The three-year project, "Akie in Tanzania – documenting a critically endangered language" (AZ. 86 405), is directed by Karsten Legère and Christa König.

References

Bakken, M.H. 2004. Becoming Visible: Economic and Social Transformation and Marginalization of Akie Hunters and Gatherers in Northern Tanzania. PhD dissertation, University of Oslo.

Batibo, H. 1992. The fate of ethnic languages in Tanzania. In Brenzinger (ed.), 85–98.

Blench, R. 2007. Endangered languages in West Africa. In Brenzinger (ed.), 140–162.

Brenzinger, M. (ed.). 1992. *Language Death. Factual and Theoretical Explorations with Special Reference to East Africa* [Contributions to the Sociology of Language 64]. Berlin: Mouton de Gruyter. doi:10.1515/9783110870602

Brenzinger, M. 2003. Language and conceptual diversity under threat: Language endangerment on the African continent. In *Linguistic Cultural Identity and International Communication*, J. Vielberth & G. Drexler (eds), 59–77. Saarbrücken: AQ-Verlag,

Brenzinger, M. 2007a. Language endangerment in Southern and Eastern Africa. In Brenzinger (ed.), 179–204. doi:10.1515/9783110197129

Brenzinger, M. (ed.). 2007b. *Language Diversity Endangered* [Trends in Linguistics, Studies and Monographs 181]. Berlin: Mouton de Gruyter. doi:10.1515/9783110197129

Greenberg, J.H. 1963. *The Languages of Africa*. The Hague: Mouton.

Guthrie, M. 1948. *The Classification of the Bantu Languages*. London: International African Institute.

Heine, B. 1980. *The Non-Bantu Languages of Kenya* [Language and Dialect Atlas of Kenya 2]. Berlin: Dietrich Reimer.

Heine, B., König, C. & Legère, K. 2014. What does it mean to be an endangered language? The state of Akie, a Tanzanian language. In *Current Research in African Studies: Papers in Honour of Mwalimu Dr. Eugeniusz Rzewuski*, I. Kraska-Szlenk & B. Wójtowicz (eds), 107–122. Warsaw: Dom Wydawniczy Elipsa.

Kaare, B.T.M. 1996. The Symbolic Construction of Community Identity of the Akie Hunter-gatherers of Northern Tanzania. PhD dissertation, London School of Economics and Political Science.

Kenny, M.G. 1981. Mirror in the forest: The Dorobo hunter-gatherers as an image of the other. *Africa* 15(1): 477–495. doi:10.2307/1158950

Klumpp, D. & Kratz, C. 1993. Aesthetics, expertise, and ethnicity: Okiek and Maasai perspectives on personal ornament. In *Being Maasai: Ethnicity & Identity in East Africa*, T. Spear & R. Waller (eds), 195–221. Oxford: Currey; Dar es Salaam: Mkuki na Nyota; Nairobi: EAEP; Athens: Ohio University Press.

König, C., Heine, B. & Legère, K. 2014. *The Akie language of Tanzania: Observations on their language and culture*. Typescript, Universität zu Köln/Universität Wien.

König, C., Heine, B. & Legère, K. 2015. *The Akie Language of Tanzania: A Sketch of Discourse Grammar*. Tokyo: Research Institute for Languages and Cultures of Asia and Africa (Tokyo University of Foreign Studies).

Krauss, M. 2007. Classification and terminology for degrees of language endangerment. In Brenzinger (ed.), 1–8.

Legère, K. 1992. Language shift in Tanzania. In Brenzinger (ed.), 99–115.

Legère, K. 2006. Language endangerment in Tanzania: Identifying and maintaining endangered languages. *South African Journal of African Languages* 26(3): 99–112.

Legère, K. 2008. Vidunda (G38) as an endangered language? In *Vanishing Voices: The Endangered Languages across the Globe*, S. Priya (ed.), 83–100. Bangalore: Icfai Research Center.

Legère, K. 2012. Endangered languages in Africa: Focus on Tanzania's Ngasa and Akie. In *Issues of Language Endangerment*, X. Shixuan, T. de Graaf & C. Brassett (eds), 89–102. Beijing: Chinese Academy of Sciences.

Maghimbi, S. 2005. The dynamics of household livelihood strategies in the drylands of Central and North Eastern Tanzania: The case of Kiteto District. In *Biodiversity Research for Livelihood Support and Food Security*, B. Kamondo et al (eds), 145–149. Nairobi: National Museum of Kenya.

Maguire, R.A.J. 1948. Il-Torobo, Being some notes on the various types of Dorobo found in the Masai reserve of Tanganyika territory and contiguous districts. *Tanganyika Notes and Records* 26: 1–28.

MLT (Mradi wa lugha za Tanzania [Languages of Tanzania Project, LoT]) 2009. *Atlasi ya Lugha za Tanzania* (Language Atlas of Tanzania). Dar es Salaam: University of Dar es Salaam.

Nurse, D. & Philippson, G. (eds). 2003. *The Bantu Languages*. London: Routledge.

Petrollino, S. & Mous, M. 2010. Recollecting words and phrases in Aasá, a dead language in Tanzania. *Anthropological Linguistics* 52(2): 206–216. doi:10.1353/anl.2010.0012

Rottland, F. 1982. *Die südnilotischen Sprachen: Beschreibung, Vergleichung und Rekonstruktion* [Kölner Beiträge zur Afrikanistik 7]. Berlin: Reimer.

Rottland, F. & Voßen, R. 1977. Grundlagen für eine Klärung des Dorobo-Problems. In *Zur Sprachgeschichte und Ethnohistorie in Afrika*, W.J.G. Möhlig, F. Rottland & B. Heine (eds), 213–238. Berlin: Reimer.

Schöperle, F. 2011. The Economics of Akie Identity: Adaptation and Change among a Hunter-gatherer People in Tanzania. MA thesis, University of Leiden.

Sommer, G. 1992. A survey on language death in Africa. In Brenzinger (ed.), 301–417.

UNESCO Ad Hoc Expert Group on Endangered Languages. 2003. Language Vitality and Endangerment. Document submitted to the International Expert Meeting on UNESCO Programme Safeguarding of Endangered Languages. *Paris*, 10–12 March.

Language transmission and use in a bilingual setting in rural Tanzania

Findings from an in-depth study of Ngoni

Tove Rosendal
University of Gothenburg

When studying languages in a bi- or multilingual language contact situation, it is necessary to go beyond the schematic evaluations and scales describing language endangerment which have been developed over the last decades. Here the case of the Tanzanian language Ngoni, represented by quantitative sociolinguistic data from fieldwork in a rural area in the Ruvuma Region in Tanzania, shows that a too general assessment can be misleading. In Tanzania it is the African language Swahili, and not the global language English and ongoing globalization, which at present represents the major threat to other African languages and the maintenance of these languages. The study indicates that Ngoni is more at risk than would have been judged from endangerment scales.

Keywords: African language, globalization, language contact, language transmission, maintenance, Ngoni, Swahili, Tanzania

1. Introduction

When studying languages in a language contact situation it is necessary to go beyond the schematic evaluations and scales describing language endangerment which have been developed over the last decades, often with the ambition to revitalize languages. The work of Fishman, for example, Fishman's GIDS (1991), subsequently developed and refined by Brenzinger et al. (2003), Lewis and Simons (2010), and Moseley and Nicholas (2010), among others, provides a valuable metric tool for assessing the status and vitality of languages. These scales use assessments, estimations and evaluations of factors such as intergenerational transmission of language, number of speakers, loss of domains, response to new domains and media, materials for education and literacy, government and institutional attitudes, community member attitudes, and the amount of quality

documentation of the language under scrutiny. Other factors, in addition to these, may be important to include for understanding processes of language endangerment and loss.

The Ethnologue (Lewis, Simons & Fennig 2015) stresses that the complexity of interrelated factors requires a new schema for evaluation of endangerment, and promotes the Expanded Graded Intergenerational Disruption Scale (EGIDS) with 13 levels, developed by Lewis and Simons in 2010. Moseley (Abstract LAUD 2014), for example, additionally investigates whether vernacular literacy is correlated to language endangerment. In fact, no single factor should be studied in isolation, as stated by Brenzinger et al. (2003). All these external and social factors that compose the extralinguistic environment must be taken into account to evaluate the status of a language in a bilingual contact situation with a more prestigious language, even if some factors, e.g. intergenerational language transmission, are considered especially important for language maintenance (Austin 2008; Grenoble 2011; Legère 2007; Norris 2010). However, evaluations using these scales are often not based on in-depth studies, but rather give a general or, at times, subjective overview of the situation. Furthermore, some of these factors are not easily obtained from statistics and general data, as pointed out by Lewis (2005: 27). Attitudes towards the L1 and intergenerational language transmission are such factors which are hard to evaluate without conducting systematic fieldwork. Although language contact is a prerequisite for language endangerment, language contact does not always lead to language loss. Only a more fine-grained study, adapted to the local context, can provide an evaluation of language endangerment and the necessary understanding of processes at work regarding language maintenance or loss.

The study of the Tanzanian language Ngoni, which is presented here, summarizes the results of an extensive survey of sociolinguistic conditions in rural Tanzania. It goes beyond the more easily obtained information about extralinguistic factors which influence language use in a contact situation. By primarily focusing on the sociolinguistic setting, language preferences, intergenerational language transmission and community members' attitudes (here represented by how children assess the use of their L1 and the status language Swahili), the study contributes to understanding not only how Ngoni, but most Tanzanian languages, are affected by the long and extensive promotion of the national language Swahili.[1]

Based on 784 interviews following a questionnaire the study provides a solid basis for evaluating the viability of a Tanzanian language which, seen from the number of speakers only, at first sight would not even have been considered for

---

1.   All languages in Tanzania are probably in a rather similar situation. However, each language has to be investigated based on its specific sociolinguistic and socio-demographic background.

an evaluation regarding endangerment. This more in-depth study of language vitality may be labelled a grounded endangerment evaluation, as opposed to more general assessments.

First, an overview of the linguistic situation and the background of the study will be given, followed by a discussion of the theoretical assumptions the study is based on. Subsequently, the study design and methods are described. The findings and their implications are then discussed and are in the final section evaluated from a language endangerment point of view.

2.   Background and language situation

In Tanzania it is the African language Swahili, and not the global language English and ongoing globalization, which at present represents the major threat to linguistic diversity, as also stressed by Heine (this volume). Swahili in Tanzania is affecting language ecology in a different way than it, for example, is causing language shift in urban contexts in Kenya and DR Congo, as described by Gibson and Araali (this volume). Swahili is the national and co-official language of Tanzania, alongside English, and is used within education and other formal domains and as a language of wider communication. The Tanzanian Bantu language Ngoni, N12, (Guthrie 1948, 1971) is one of more than 150 languages or dialects in Tanzania (Muzale & Rugemalira 2008: 80). For many years, Ngoni and other Tanzanian languages have been discouraged due to fear of tribalism. Although the local ethnic languages were mentioned in the *Sera ya Utamaduni* (Tanzania 1997), the cultural policy document where the Tanzanian official position regarding national languages other than Swahili is summarized, nothing has been done to introduce them for use in any formal domains of society. This policy has undermined both the status and use of these languages.

Ngoni is spoken in the Ruvuma Region in the Southern Highlands of Tanzania, and has 258,218 speakers (LOT 2009). Although the number of speakers is relatively high, the sociolinguistic situation with Swahili as the more prestigious language and as a national and widely used lingua franca poses a serious threat to the Ngoni language and culture.

The data of this case study of Ngoni are based on empirical fieldwork conducted in six villages north-west of Songea in Tanzania in November–December 2012.[2] See Figure 1 for a map of the Ngoni-speaking area.

Figure 1. Map of the Ngoni speaking area

The area is homogeneous with predominantly Ngoni speakers, but as the study shows, other Tanzanian languages are used, especially in some villages. In such an environment it would be expected that a child would learn to speak and understand the dominant L1 of the area without major difficulties, provided that ethnic identity is strong and that intergenerational language transmission functions without disruptions. Therefore, the interviews which followed a questionnaire, focused, among other things, on:

- when and how these children acquired their first language (L1) and the lingua franca Swahili;
- language preferences;
- how well they themselves estimated their mastery of the languages, in addition to factors such as
- family background (origin of mother and father and their first languages);
- attitudes towards the language(s) used in this environment.

The above factors are some of the important factors which maintain cultural identity in general, and language maintenance and vitality in multilingual countries in particular (Batibo 2005; Brenzinger 1992, 1998, 2007; Calvet 1987; Crystal 2000; Fishman 2001).

Two previous studies conducted in Kilagano Ward, Songea Rural District, in 2008 and 2011 observed that adult speakers of Ngoni have negative attitudes towards their L1 and favour Swahili (Mapunda 2013). This attitudinal orientation was also noted by Ngonyani (2003: 4) among young Ngoni, and in a recent study on lexical borrowing from Swahili (Rosendal & Mapunda 2014), which included interviews with informants to provide metadata for a study on lexical borrowing.

The traditional diglossia situation (Fishman 1967, 1970), where a high-status code or language, like Swahili, is used in prestigious domains and low-status languages like Ngoni are used in non-formal domains, is changing.[3] Interviews which were conducted with informants to provide metadata for the study of borrowing clearly showed that the use of Swahili was frequent in activities outside the family/home and that Ngoni often was mixed with Swahili within the family.

In a stable language situation with no threat of language shift, i.e. a "gradual displacement of one language by another in the lives of community members" (Dorian 1982: 44), it would be expected that a child living in the countryside in the Ngoni area would learn to speak Ngoni if both parents are Ngoni. Also, traditional

---

3.   In fact, in Tanzania as in many African countries there is a conflictual triglossic environment with the ex-colonial language (English) at the top of the hierarchy, a regional lingua franca, here Swahili, in the middle, and the local community language at the bottom of the heap (Rugemalira 2013).

stories and songs would be expected to be taught to children and learnt by minority groups in the area. However, macro factors such as intensity of contact with other languages and cultures, cultural pressure and language attitudes influence language use and play an important role in determining how well languages and cultures are preserved. With this as a point of departure, the present sociolinguistic study of school children provides further insight into the vitality of the Ngoni language and culture. Aspects such as language preferences and transmission will be discussed in some detail, based on the survey, while other questionnaire data will only be summarized as a background for the discussion.

Based on previous studies and knowledge of the sociolinguistic situation, Swahili was expected to be found to be the high-status language used frequently in the local community and within family life. This language use was furthermore expected to be reflected in language preferences and attitudes, which in turn would be reflected in lack of knowledge of Ngoni culture. Before discussing the results, a summary of the methodology used for the study is given.

3.   Data collection and methods

This paper is based on empirical fieldwork and results from interviews with 784 primary school children from six village schools (grades 1, 2, 5 and 6) in a rural area northwest of Songea in the Ruvuma Region in southwestern Tanzania, about 1,000 kilometres from Dar es Salaam and the coast. In this region, the Bantu language Ngoni is traditionally the main first language (L1).

The study was conducted in the Songea Rural District, east of Lake Nyasa in November–December 2012. Songea Rural District is situated geographically in the middle of the Ngoni-speaking area. The data were collected in the six small villages Mgazini, Kilagano, Mhepai, Lugagara, Muungano Zomba and Lihanje, situated in an area 55–80 kilometres northwest of Songea (see Figure 1). These villages are small and quite isolated. People living in the area are mostly subsistence farmers.

All pupils in the targeted grades who were present the day the study was conducted were interviewed following a questionnaire with 17 questions. As already mentioned, the study comprised 784 pupils in total, ranging from 95 to 147 pupils per school. 52.8 per cent of the interviewed pupils were grade 1 and 2 pupils, while 47.2 per cent were in grades 5 and 6. In total, 45.4 per cent were male, while 54.6 per cent were female. Table 1 gives an overview of the schools and the number of interviewees.

The interviews were conducted by the researchers with the assistance of local staff. Grade 1 and grade 2 pupils were interviewed individually. Grade 5 and 6

Table 1. Primary schools and number of interviewed children

School	Grade 1	Grade 2	Grade 5	Grade 6	Total number of interviews/school
Mgazini	29	27	24	15	95
Lihanje	43	31	32	36	142
Mhepai	32	26	36	25	119
Kilagano	35	35	40	38	148
Lugagara	59	22	32	20	133
Zomba	44	32	41	30	147
TOTAL	243	173	205	164	784

pupils filled in the questionnaire in the classroom, with considerable assistance from the researchers and trained assistants. Four researchers and 2–4 local helpers participated in each school. Prior to the data collection thorough training was conducted, where the aims of the study, all of the questionnaire questions and interpretations of the questions as well as possible answers were meticulously explained and discussed. Most students of grades 5 and 6 in these schools needed extensive help, not only with reading and writing, but also regarding understanding the questions. When the study was carried out, one of the researchers therefore explained each question in detail and the different answer alternatives were exemplified orally and with the help of the blackboard. The pupils were also assisted in writing the answers – a time-consuming method but imperative to get useful and reliable data. The explanations and the interviews were mainly conducted in Swahili, but when there were communication problems, it was possible to switch students among the Ngoni-speaking researchers/assistants and these students were interviewed in Ngoni.

Two programs were used to analyse the data: SPSS and SAS. SPSS was used for data entry, validation and cleaning and for analyzing the multiple response questions and presentation of cross-tabulations. SAS was used to show the frequency distribution of the variables used in the study.

4. Findings and their implications

The data from the interviews in the six village schools showed that Ngoni was spoken by 91.6 per cent of the pupils. Other languages spoken were (in descending order) Matengo (5.4%), Bena and Swahili (0.9%), Ndendeule and Pangwa (0.4%), Yao (0.3%) and Kinga (0.1%), as shown in Table 2 which gives a cross-tabulation of L1 and language preferences.

Table 2. Cross-tabulation of L1 and language preferences

| Language | Preferred language | | | | Reported L1 |
	Frequency	Per cent	Cumulative frequency	Cumulative per cent	Per cent
Swahili	281	36.21	281	36.21	0.92
Ngoni	457	58.89	738	95.10	91.55
Matengo	29	3.74	767	98.84	5.42
Ndendeule	1	0.13	768	98.97	0.40
Bena	3	0.39	771	99.36	0.92
Yao	2	0.26	773	99.61	0.26
Nyasa	1	0.13	774	99.74	–
Sukuma	1	0.13	775	99.87	–
Pangwa	1	0.13	776	100.0	0.40
Kinga	–	0.00	776	100.0	0.13
Frequency missing = 8					

As can be seen in Table 2, the interviewees' language preferences do not reflect their first languages. Surprisingly, a total of 36.2 per cent of the children claimed they prefer to speak Swahili and not their L1. Only 58.9 per cent of the children said that they prefer to speak Ngoni and 4.9 per cent the remaining L1s, in spite of the fact that these languages are the L1s of 99.1 per cent of the interviewees.

As Swahili is the medium of instruction from grade 1 onwards, information about when and where the children learnt Swahili is of interest. The data showed that of those who preferred to speak Swahili and not their L1, nearly 65 per cent had learnt the language before entering school – and 35 per cent after. Swahili is thus clearly present in the daily life of children in the area, outside formal activities.

It would be natural to find that some children, coming from families where either the mother or father are from different ethnic groups, use Swahili more at home and prefer Swahili. However, the data revealed that as much as 32.4 per cent of the children with both a Ngoni mother and father preferred to speak Swahili and not their own L1. The data from the questionnaires furthermore showed that 81.8 per cent of the children claimed that they speak Ngoni with ease. Thus, language preference is for most of these children not linked to lack of knowledge of the language but to other factors such as prestige, attitudes and communication structure. In the following, some aspects of language use will be discussed in more detail.

The study showed that the children mainly spoke their L1 with relatives (guardians, grandparents and siblings were the most frequent categories) and friends and that they used the L1 at home (93.3%). Generally speaking, girls more

Table 3. Children whose parents are both Ngoni: Cross-tabulation of when Swahili is learnt and language preference

Time Swahili was learnt	Preferred language					
	Swahili	Ngoni	Matengo	Ndendeule	Sukuma	Total
Before starting school						
Frequency	118	160	1	1	1	281
Per cent	19.90	26.85	0.17	0.17	0.17	47.15
After starting school						
Frequency	75	239	1	0	0	315
Per cent	12.58	40.10	0.17	0.00	0.00	52.85
Total						
Frequency	193	399	2	1	1	596
Per cent	32.38	66.95	0.34	0.17	0.17	100.00
Frequency missing = 9						

often acquired the language at home, while boys more often learnt the L1 in the street, from friends before school age and at school.

In total, a little more than half of the informants (50.5%) learnt Swahili before starting school. Additionally, more children who said they prefer to speak Swahili and not an L1 learnt Swahili before starting school. As would be expected, more children with one parent who was not of Ngoni origin started to speak Swahili before entering school, compared to children for whom both parents were Ngoni: 60 per cent when the father was not Ngoni and 71.4 per cent when the mother was not Ngoni. As mentioned earlier, this should be compared to a value of 47.1 per cent when both parents were Ngoni. Still, this is a high figure which shows that Swahili is learnt at an early age even when there probably are no practical needs for the language as a lingua franca within the family. The data confirmed that Swahili even in this rural area is a high status language which exists alongside the L1 in the local communities and within the family.

If the answers of the 32.4 per cent of the respondents who came from purely Ngoni families (where both parents were ethnically Ngoni) and who preferred to speak Swahili are studied, 19.8 per cent had learnt Swahili before starting school and 12.6 per cent learnt Swahili after entering school. Of the total number of interviewees with a Ngoni ethnic background, nearly 67 per cent preferred Ngoni. Only a few (a total of 4 persons) stated other Tanzanian languages such as Matengo, Ndendeule and Sukuma. Table 3 gives the results in more detail.

These data confirm that Swahili is spoken even in linguistically homogeneous families and that Swahili over time has come into daily use, which was also found in Rosendal and Mapunda (2014). The latter study additionally showed that

people code-switch between Swahili and Ngoni, even when prompted to only speak Ngoni.

Generally speaking, a rather low value seems to be attached to the L1s in the area, especially to Ngoni. At the same time, the results of the study demonstrate the prestige attached to Swahili, which was claimed to be the L1 of less than 1 per cent of the informants. Furthermore, this sociolinguistic situation affects language proficiency. Although the majority of the Ngoni children claimed to speak Ngoni with ease, nearly 20 per cent (18.2%) understood or spoke their L1 with difficulty or with great difficulty.

Since attitude is an important factor when discussing language maintenance, questions regarding attitudes towards the L1 were highlighted in the questionnaire – with dichotomized yes/no questions and an open question where the children could give personal and individual answers as to why (or why not) the L1 should be used. The following trends were noted:

28.9 per cent of the respondents did not find it important to use an L1 such as Ngoni. There were no statistically significant differences between the younger children (grades 1 and 2) and the older respondents (grades 5 and 6).

Results to an open question about why or why not the L1 should be used were sorted into major categories, as demonstrated in Table 4. The major advantages of using the L1 may be grouped into 'communicative purposes' (accounting for a total of 64.1% of responses), 'cultural reasons' (17.7%) and 'emotional motives' (11.9%).

The major disadvantages of using mother tongues were said to be the lack of communicative function and usefulness (accounting for a total of 65.6% of

Table 4. Advantages of using the L1 (open question)

(Q17) Advantages of using L1	Responses	
	N	%
Makes me able to speak with and understand easily family and other people who speak the same language	433	58.9%
L1 is the most spoken language	30	4.1%
I like it, it's my L1	85	11.6%
You interact more easily than with Swahili	38	5.2%
Helps to understand and learn from past events	60	8.2%
To maintain my L1, culture and tradition	52	7.1%
Make yourself happy	2	0.3%
Identity of your tribe	18	2.4%
Other	17	2.3%
TOTAL	735	100.0%

Table 5. Disadvantages of using L1 (open question)

(Q17) Disadvantages of using L1	Responses	
	N	%
Other people will not understand me if I speak the L1	43	25.6%
Swahili is used at school and by many people/L1 not used at school	29	17.3%
Parents of a different tribe don't speak L1	1	0.6%
People laugh at us if we speak L1	5	3.0%
I don't know it well/it is tough	24	14.3%
I don't like it	17	10.1%
Has no importance/not a national language	12	7.1%
I like Swahili/Swahili is a national language	28	16.7%
Other	9	5.4%
TOTAL	168	100.0%

responses) and negative and emotional attitudes towards the L1 (31.5%), as demonstrated in Table 5.

As can be seen in the above data, more children listed advantages than disadvantages in using their L1. Although 17.7 per cent of the listed positive answers may be linked to culture preservation (to maintain my L1, culture and tradition (7.1%), identity of your tribe (2.4%), help to understand and learn from past events (8.2%)), the figures for knowledge of Ngoni culture, which was included in the study, were quite low.

As many as 46 per cent of all children whose mother and father were both of Ngoni origin did not know any traditional Ngoni stories. The equivalent figures for knowledge of Ngoni songs were 67.2 per cent (yes) and 32.8 per cent (no). Thus, more of the Ngoni children knew songs than stories. Furthermore, the data showed that girls knew both stories and songs more frequently than boys.

Considerably fewer children whose L1 was not Ngoni knew any Ngoni stories (24.7%) or Ngoni songs (29%). The number of respondents was, however, quite low, but nevertheless indicates that Ngoni does not have the impact one would expect a majority culture/language to have on small minorities living in such a rather homogeneous rural environment. The question which arises is: Why do not *all* children know these cultural expressions in such an environment? Why is this knowledge not transferred to the children in these ethnically rather homogeneous communities? Is it because not even the parents' generation know them or use them any more – and that the common attitude among the Ngoni is that the Ngoni culture, including language, is not important or even not valuable?

5.   Vitality of Ngoni vs. endangerment scales

It is clear that Ngoni is under pressure due to the influence of Swahili, which has a status that Ngoni does not have. The fact that one third of the children living in an ethnic Ngoni community, with a Ngoni mother and father and Ngoni as their L1, prefer to speak Swahili is a clear indication of the low esteem attributed to Ngoni. As a consequence, Swahili is learnt at an early age even when there probably are no practical needs for the language as a lingua franca within the family. Additionally, Ngoni culture is not valued, as seen from the data. The high regard attached to Swahili was also reflected in the data about attitudes among the children towards Ngoni and other L1s. Even though the children value their L1 as a medium of communication with family and friends, it is nevertheless clear that they do not see Ngoni as having any important function in the local Tanzanian setting. However, it is possible that the children to some degree hide relevant information, due to the strict medium-of-education policy, which discredits L1 use and does not tolerate any languages but Swahili. As a result of this education policy, even the process of L1 consolidation among young children may be affected. The data showed reduced language proficiency. This indicates that the future of Ngoni language and culture is under threat. The data from the fieldwork furthermore demonstrates that Ngoni identity is rather weak.

The results of the study, combined with the previous studies of Ngoni in a language contact situation, support the assumption that Ngoni is under substantial pressure. Although the number of speakers is quite high, Ngoni is not used in any domains outside non-formal communication, neither used in media nor education. Furthermore, Ngoni, like all other Tanzanian languages, has no official status and is not well documented. All these factors, together with the negative attitudes which are attached to the language, indicate that Ngoni faces severe challenges in this contact situation with Swahili, a contact situation which triggers language shift and represents a threat to linguistic diversity in the area and in Tanzania in general. Using the UNESCO degrees of endangerment (Brenzinger et al. 2003), the degree of endangerment regarding intergenerational language transmission should, based on the questionnaire results, be labelled 'unsafe' (grade 4 of 5 possible grades) and the language assessed as being used in 'dwindling domains'. According to Moseley and Nicholas' assignment criteria, Ngoni would be classified as 'vulnerable' (Moseley & Nicholas 2010).

However, it would not have been possible to evaluate intergenerational language transmission or attitudes without some kind of in-depth study, such as the one summarized in this paper. Without this more nuanced picture, language transmission would probably have been assessed and labelled as 'safe' (grade 5 of 5) – because the language is used by all ages, from children up. Ethnologue

(Lewis, Simons & Fennig 2015), for example, judges Ngoni as level 5, 'developing', which implies that "the language is in vigorous use, with literature in a standardized form being used by some though this is not yet widespread or sustainable". This grading is clearly above the endangerment side of the EGIDS scale. The more comprehensive set-up of this study thus leads to a more negative conclusion. Furthermore, it is not possible to seriously evaluate attitudes without detailed and extensive fieldwork. Without a grounded knowledge of the local Ngoni community, it would probably be assumed that all members in such a homogeneous rural setting would value their language and wish to see it promoted (i.e. grade 5 of 5 according to the UNESCO scale). The study clearly showed that the attitudes among children in this community are quite different. Thus, endangerment scales must at least be complemented with more grounded fieldwork to be reliable.

This study has contributed to the understanding of the language shift question in giving a more thorough, detailed and nuanced investigation of some of the most important factors which are used to indicate language endangerment, especially attitudes towards the L1 and how children acquire their L1 and Swahili. However, the mechanisms which influence language use, language changes, attitudes and status in the rural communities in the Songea District need further investigation. A natural next step would be to study code-switching practices linked to identity and whether or not the loss of the Ngoni language and culture is a sign of loss of the Ngoni identity.

The Tanzanian language policy, which promoted an African and indigenous language as the official language at independence, has often been praised as a bold and better choice than maintaining ex-colonial languages such as English and French. However, the language situation in Tanzania shows similar trends to what we today see in an increasingly globalized world where English represents a threat to linguistic diversity. There is a major difference though, and this difference lies in the successful implementation of Swahili and the comprehensive influence and impact the language today has on language practices and preferences all over Tanzania and, as we have seen, even in the most remote and rural areas. Furthermore, the result of the Tanzanian policy, which aimed at giving all Tanzanians equal opportunities through Swahili as a medium of instruction, has probably achieved the opposite: the policy led to a marginalization of most of the rural population as their Swahili knowledge is not good enough to provide them educational success.

Although the influence of Swahili is possibly even stronger in the coastal areas of eastern Tanzania, the data from the village school survey in the Ruvuma Region in southwestern Tanzania indicate that the same processes are at work in this rural setting. To counteract language loss, which is the result, official policy must recognize the value of all Tanzanian languages. So far, this is unfortunately not on the agenda, even if lip service occasionally is paid.

References

Abstracts LAUD 2014. 36th International LAUD Symposium. March 31–April 3, 2014. Landau: University of Koblenz-Landau.

Austin, P.K. (ed.). 2008. *One Thousand Languages: Living, Endangered and Lost*. Berkeley CA: University of California Press

Batibo, H. 2005. *Language Decline and Death in Africa*. Clevedon: Multilingual Matters.

Brenzinger, M. 1992. *Language Death: Factual and Theoretical Explorations with Special Reference to East Africa*. Berlin: Mouton de Gruyter. doi:10.1515/9783110870602

Brenzinger, M. 1998. *Endangered Languages in Africa*. Cologne: Rüdiger Köppe.

Brenzinger, M. (ed.). 2007. *Language Diversity Endangered*. Berlin: Mouton de Gruyter. doi:10.1515/9783110197129

Brenzinger, M., Dwyer, A.M., de Graaf, T., Grinevald, C., Krauss, M., Miyaoka, O., Ostler, N., Sakiyama, O., Villalón, M.E., Yamamoto, Y. & Zapeda, O. (eds). 2003. *Language Vitality and Endangerment*. Paris: UNESCO.

Calvet, L.-J. 1987. *La guerre des langues et les politiques linguistiques*. Paris: Payot.

Crystal, D. 2000. *Language Death*. Cambridge: CUP. doi:10.1017/CBO9781139106856

Dorian, N.C. 1982. Language loss and maintenance in language contact situations. In *The Loss of Language Skills*, L. Lambaret & B. Freed (eds), 44–59. Rowley MA: Newbury House.

Fishman, J.A. 1967. Bilingualism with or without diglossia; diglossia with and without bilingualism. *Journal of Social Issues* 23: 29–38. doi:10.1111/j.1540-4560.1967.tb00573.x

Fishman, J.A. 1970. *Sociolinguistics*. Rowley MA: Newbury House.

Fishman, J.A. 1991. *Reversing Language Shift: Theoretical and Empirical Foundations of Assistance to Threatened Languages*. Clevedon: Multilingual Matters.

Fishman, J.A. 2001. *Can Threatened Languages Be Saved: Reversing Language Shift*, Revisited: A 21st Century Perspective. Clevedon: Multilingual Matters.

Grenoble, L. 2011. Language ecology and endangerment. In *The Cambridge Handbook of Endangered Languages*, P.K. Austin & J. Sallabank (eds), 27–44. Cambridge: CUP. doi:10.1017/CBO9780511975981.002

Guthrie, M. 1948. *The Classification of the Bantu Languages*. London: OUP for the International African Institute.

Guthrie, M. 1971. *Comparative Bantu*. Farnborough: Gregg International Publishers.

Legère, K. 2007. *Vidunda (G38) as an endangered language?* Somerville MA: Cascadilla Procedings Project.

Lewis, M.P., Simons, G.F & Fennig, C.D. (eds). 2015. *Ethnologue: Languages of the World*, 18th edn. Dallas TX: SIL International.

Lewis, M.P. & Simons, G.F. 2010. Assessing endangerment: Expanding Fishman's GIDS. *Romanian Review of Linguistics* 55(2): 103–120.

Lewis, M.P. 2005. Towards a categorization of endangerment of the world's languages. <http://www-01.sil.org/~simonsg/preprint/EGIDS.pdf>

Loh, J. & Harmon, D. 2014. Comparing status and trends in linguistic and biological diversity. *Abstracts* LAUD 2014, 91–93.

LOT. 2009. *Languages of Tanzania Project. Atlasi ya Lugha za Tanzania*. Dar es Salaam: LOT.

Mapunda, G. 2013. Ngoni people's attitudes towards the use of Kingoni in beginner classes. *Journal of Linguistics and Language in Education* 7(1): 74–92.

Moseley, Ch. 2014. Language endangerment and vernacular literacy. *Abstracts* LAUD 2014, 102–103.

Moseley, Ch. & Nicholas, A. 2010. *Atlas of the World's Languages in Danger*. Paris: UNESCO Publishing.

Muzale, H.R.M. & Rugemalira, J.M. 2008. Researching and documenting the languages of Tanzania. *Languages, Documentation and Conservation* 2(1): 68–108.

Ngonyani, D. 2003. *A Grammar of Chingoni*. Munich: Lincom.

Norris, M.J. 2010. Canada and Greenland. In *Atlas of the World's Languages in Danger*, C. Moseley (ed.). Paris: UNESCO Publications.

Rosendal, T. & Mapunda, G. 2014. Is the Tanzanian Ngoni language threatened? A survey of lexical borrowing from Swahili. *Journal of Multilingual and Multicultural Development* 35(3): 271–288. doi:10.1080/01434632.2013.864659

Rugemalira, J. 2013. The Tanzania experience in language policy and planning. In *Language Policy in Africa: Perspectives for Cameroon*, P. Akumbu & B. Chiatoh (eds), 62–97. Kansas City KS: Miraclaire Academic Publications.

Tanzania, United Republic of. 1997. *The Cultural Policy/Sera ya Utamaduni*. Dar es Salaam: Mpiga Chapa Mkuu wa Serikali.

Language shift and endangerment in urban and rural East Africa

Three case studies

Maik Gibson*, ** and B. Araali Bagamba**

* Redcliffe College / ** SIL International

In this paper we present empirical data on three places in East Africa where language shift is occurring, two urban and one rural. In doing so, we hope that both the similarities and the differences will illuminate some of the underlying factors which influence communal choices about language maintenance and shift. All three case studies are drawn from our own work, and one is part of a larger ongoing project concerning urban language shift in East Africa. Bagamba and Gibson (forthcoming) reports on one other context in the same project, that of the Kisii community in Homa Bay, Kenya.

Keywords: East Africa, choices, ecology, EGIDS, maintenance, Nairobi, shift, Swahili, urban/urbanisation

0. Introduction

African cities are often relatively recent foundations (e.g. Nairobi dates from 1899), and as such are sites of recently developed and dynamic language ecologies; the question of what commonalities these ecologies share is part of the rationale behind this paper. And yet we also ask the question, which is yet to be answered in a coherent way, of what influence the recently emerged urban language ecologies have on the rural contexts that surround them. It is certain that different dynamics are at work in the city and the village, and yet these dynamics do not exist in isolation from each other – people, and with them their language practices, move from one to the other daily.

These are by no means the first studies of language shift in East Africa. The contributions in Brenzinger (1992) cover various cases, both in terms of general trends, for example in Tanzania, and specific, mainly advanced cases of language

shift, such as among the Suba and Dahalo of Kenya. While urban shift is mentioned by Legère (1992:100) for Mbeya, the majority of cases concern shift in rural populations. Most of these cases of shift (see Batibo 1992:92 for a list) are from one ethnically-marked language to another, alongside other (including urban) cases where the shift is to the lingua franca Swahili, which away from the coast does not correspond to an ethnic shift.

The contexts we look at are (1) urban Nairobi (Gibson 2012), in particular in the informal settlement of Kibera, (2) the Bhadha-speaking Northern Hema in the town of Bunia, Eastern Democratic Republic of Congo (DRC), from unpublished research undertaken by both authors, and (3) the rural homeland of the Bhele community, also in the Eastern DRC, from Bagamba's own unpublished research. In each of these cases, the target of the shift is a local variety of Swahili, which is not a language that carries an ethnic identity in any of these contexts. We do not claim that any of these studies represents the final word on what is happening in these communities.

With Fishman's claim (1971:315) that "urban dwellers are more inclined to shift; rural dwellers are less inclined to shift", we might assume that African cities are sites of widespread language shift. However, Mufwene (2010:915) questions whether this applies in all African contexts, suggesting that at least parts of cities may be seen as "mega-villages", with a mix of rural and urban dynamics. Here, as people from one ethnic community may well choose to live in the same part of a city as each other, we can envisage certain parts of a city being environments which favour ethnic language maintenance, along with the practice of an urban language of wider communication (LWC). To use metaphors developed by Gary Simons, such a city with mega-villages would be framed as a salad bowl, where each ingredient maintains its distinctiveness, as opposed to the more commonly used metaphor of the melting pot, where different ingredients come together to form something new and relatively uniform – this would represent a case where different communities all shift to an LWC. The first study addresses the case of Nairobi, asking the question whether it may be best seen as a melting pot (where all shift is to Swahili[1] and/or English, while losing competence in the communal languages), or as a salad bowl, where there is practice of LWC(s) but also of communal languages being maintained in different neighbourhoods. The new metaphor developed for this paper, that of the chilli pot, where the locally-dominant communal language becomes the LWC of the city, does not apply here, but seems to apply in the Homa Bay study (Bagamba & Gibson forthcoming).

---

1.   In this case a non-standard variety of Swahili, subsuming ways of talking that might also be identified by the term *Sheng*.

The study of language shift and maintenance has often been of contexts where the end result is communal monolingualism (e.g. in Europe or North America). In none of the three case studies examined here is monolingualism the pattern adopted by those who have shifted away from a communal language. Instead, new patterns of multilingualism are adopted. We are not able to claim that this is a universal pattern for African cities or villages, but the results of these three case studies did not come as a surprise to the authors – it may well be that societal multilingualism is a norm in many parts of Africa (see Lüpke & Storch 2013: 275–278, for example). We therefore need to be careful to clarify what is inferred by a statement of language shift, avoiding the assumption of monolingualism which may well be the pervasive model elsewhere. Therefore any study of language maintenance and shift in Africa needs not just to ask what is lost, but also investigate what is maintained or gained. We now turn to look at the first case study, that of Nairobi, in more detail.

1.   Language shift in Nairobi

The Nairobi study (Gibson 2012) looked at self-reported language behaviour in two contexts: from the congregation of a primarily middle-class Pentecostal church in the suburb of Buru Buru, and in the informal settlement of Kibera, which we would expect to be a place more favourable towards community language maintenance. Table 1 looks at claimed dominant language usage in the Buru Buru sample. Note that "mother tongue" in Kenyan usage refers to communal languages (i.e. neither Swahili nor English) rather than to one's dominant language.

The sample is not especially large, and we do not claim that these figures accurately represent the percentages of language use in middle-class Nairobi. Nevertheless, a strong trend of a move away from communal vernaculars as the dominant language at home is evident. This does not necessarily indicate that the individuals (most of whom had moved into Nairobi from elsewhere) have lost all competence in their communal languages, but represents a very different

Table 1. Dominant language at home, Buru Buru (N = 46) (Gibson 2012: 570)

Language	Number	Percentage
"Mother Tongue"	3	7%
Swahili	37	80%
Swahili and English	2	4%
English	4	9%

Table 2. Dominant language at home, Kibera (N = 156) (Gibson 2012: 570)

Language	Number	Percentage
Mother Tongue (MT)	109	70%
MT and Swahili	18	12%
Swahili	28	18%

ecology from most rural contexts, where ethnic-based vernaculars still dominate in the home domain. Nor, especially, should we understand stated English dominance to imply an inability to speak (non-standard) Swahili competently. All of this sample will be at the very least bilingual.

The findings in Kibera contrast starkly with Buru Buru. In Kibera the population under investigation does not inhabit carefully-planned suburbs, working in the formal economy, but an informal settlement with a much higher population density, with a mix of formal and informal employment. More details are available in Gibson (2012), but here we present the most basic trend.

The headline figure of 70% maintenance of the dominance of the communal language is strikingly different from the Buru Buru sample, and looks like it might represent the kind of figure where we might speak of the mega-village, where urban norms have not overwhelmed rural ones, but live alongside them. However most interviewees listed "mother tongue" usage only for the home domain, claiming that street interaction was primarily in Swahili. On the other hand, the interviewers did notice a lot of use of Luo, the communal language of around two-thirds of the interviewees, on the streets, and in at least one bar. Unsurprisingly, Luos showed a higher level of maintenance, and two women from other communities who had Luo husbands claimed to have added Luo to their repertoires in the home. Furthermore, a majority of other communities also claimed to have maintained their communal language in the home, so the high density of Luos in Kibera cannot be the sole factor in vernacular retention.

When interviewees were asked which was the main language they used with their children, the 70% figure declined to 31% claiming to use their mother tongue, with an additional 17% claiming to use a mix of Swahili and the communal language. And when asked to report what language their children spoke to each other, the figures declined further to 18% using mother tongue, and 12% using a mixture of mother tongue and Swahili. The remainder of 70% reported that their children used Swahili with each other; interestingly none mentioned English here, despite three interviewees saying they spoke to their children in this language. Without wider communal use, attempts to use English, the educational language, in the home domain were not adopted by the children.

Table 3. Ethnic language use at home by age of arrival in Kibera (Gibson 2012:574)

	0–9	10–19	20 and over
Luos	83% (N = 12)	67% (N = 12)	89% (N = 18)
Other groups	29% (N = 7)	77% (N = 13)	71% (N = 14)

The factor which showed the greatest difference between Luos and others was the impact of the age of arrival in Kibera upon the home language, as can be seen in Table 3.

For those non-Luos who arrived in their first decade, the trend is towards Swahili dominance, while those who arrived later mainly use their vernacular at home. However this difference of age at arrival is not apparent for the Luo community, which is consistent with some level of maintenance of Luo occurring within Kibera itself. This is perhaps unsurprising given the assumption that the majority of Kibera's population comes from this community. With a large population shift into the cities, and the fact that informal settlements such as Kibera are the point of first arrival for many of these migrants, it seems that we have a situation which fits Mufwene's notion of the mega-village, especially as there is ethnic language transmission continuing in the city itself. Without such continued migration, we might predict shift away from (especially non-Luo) ethnic vernaculars in Kibera – to be certain much more slowly than in the suburban sample. But the continued contact with and movement from the Luo homeland will presumably mean continued use of the Luo language in Kibera for the decades to come.

2.   Language shift amongst the Northern Hema in Bunia

From urban Kenya, we now turn to examine two Congolese contexts, the first of which is the repertoire of children of Northern Hema migrants into the city of Bunia in the North-East of the DRC, and parts of the Bhele community in their rural homeland around two hundred kilometres south-west of Bunia. Neither of these studies have yet been published.

Bunia is a city of around 300,000 people (Obedi 2009), which has seen rapid population growth, not just due to the rural-to-urban migration which often has economic causes, but also due to decreased security due to civil conflict, especially between 1999 and 2005. The Northern Hema shifted their vernacular from (Bantu) Oruhema (ISO 693-3 code: NIX) to (Central Sudanic) Bhadha (ISO code 693-3: LED, a language also associated with another ethnic group, the Lendu) during the twentieth century, accompanying a change in lifestyle. Further details and analysis of this shift, which is not the subject of investigation here, can be

Table 4. Competence of Northern Hema children in Bunia in Bhadha

	Sample size	Does not under-stand Bhadha	Understands but does not speak Bhadha	Speaks Bhadha competently
5–9 year olds	59	63%	27%	10%
10–13 year olds	51	27%	55%	18%
14–18 year olds	37	19%	38%	43%

found in Bagamba (2007). We know from Kaputo (1982:47) that Hema, whose homeland almost borders Bunia, had started moving to Bunia as early as 1929. They, along with the Alur, form one of the larger communities in Bunia, but do not constitute a majority.

The Bunia study is different from that in Nairobi, in that is focused on only one ethnic community. But the more fundamental difference is the methodology, which is focused on children's knowledge of the communal language Bhadha, and is based not on self reporting, but on the interviewer's assessment of the child's language level: whether the child can speak it well, understand it well, or does not understand the language at all. This was accompanied by an interview with the parents, asking them about their language practices and the reasons behind them. Note that the LWC of Bunia, as already mentioned, is Swahili, with French being used in education, and Lingala finding some uses in parts of the town.

The first thing to notice in Table 4 is that many children are learning Bhadha in this urban context, but many are not. We see both greater understanding and capacity to speak the language competently with the children who are older. We are unable to conclude whether this is due to the language often being learnt in the home after first learning Swahili, or whether the older children reflect a more robust use of Bhadha in homes a few years ago. Interestingly there is also a broadly similar trend among Kisii children in Homa Bay, Kenya (Bagamba & Gibson, forthcoming). As in Nairobi, we assume that this pattern of incomplete learning within the community would eventually lead to loss of the language in this urban environment, were it not for continued migration which might be able to sustain the language.

We were interested in whether language choice reflected different self-perceptions of identity. This is complicated, as often ethnic identity labels refer to one's origin, and an urban identity operates at a different level, in that belonging to a town is not seen as such a basic identity as a communal one, and nor is it seen as incompatible with a communal identity. Table 5 represents the children's answers to this question.

Here we see similar to patterns to those of language use – the Northern Hema identity is claimed more by the older children, and whether this is due to this

Table 5. Identity perception by the Northern Hema children living in Bunia

	Sample size	Inhabitant of Bunia	Hema living in Bunia	Northern Hema
5–9 year olds	61	74%	7%	20%
10–13 year olds	51	61%	8%	31%
14–18 year olds	38	45%	11%	45%

Table 6. Parents' language use patterns, Bunia

	Bhadha	Swahili and Bhadha	Swahili	Other combinations
Between parents	44%	27%	18%	11%
Father to child	20%	17%	59%	4%
Mother to child	21%	18%	59%	2%

identity emerging with age, or that it was more prominent a few years ago, or whether it is seen to correlate with language use, remains an open question.

While we see little difference in self-reported behaviours of either parent, we do see that Bhadha is used much less with children than with spouses. The main reason that was given for not speaking to children in Bhadha was a perceived lack of responsiveness from children when speaking the language, associated with the parents seeing little utility in teaching their children Bhadha for urban life. In addition, much child-rearing is done by older children, so children's linguistic choices will have speedy repercussions on children just a few years younger than them. There was also not a feeling among most parents that speaking Bhadha was an essential part of a Northern Hema identity. Furthermore, 16% of the mothers were not Bhadha speakers, and in general the children from exogamous marriages did not learn Bhadha.

Interestingly, another study undertaken in Bunia, Ucuon's (2012) work on the Alur, shows a very similar pattern. Again, part of the community's children has learnt the language, with older children knowing more than younger ones.

3. Language shift in the rural Bhele community

Our final case study concerns the rural Bhele (ISO 693-3 code: BHY) community, taken from Bagamba (2012). It is a small community (around 15,000) with a neighbouring community that is dominant demographically, economically and politically, the Nande. The interethnic marriage rate is around 40%, and over half of that is with the Nande. Members of the Bhele community often present themselves as Nande when outside their homeland.

Table 7. Bhele children's competence in Bhele

	Doesn't under-stand	Passive under-standing	Speaks, but not well	Speaks fluently
6–13 year-olds	25%	35%	27%	13%
14–20 year-olds	9%	15%	17%	58%

The sample under study here was taken in the community's rural homeland, rather than in a mixed urban context. The sample was from the four more accessible Bhele clans (of a total of six). The two other clans are in remote forest villages and can be reached only by a two-day journey on foot, so no conclusions about these other two clans can be made, especially as we may hypothesise that the less accessible clans may be less prone to language shift. It seems that, where found, the direction of shift is towards Nande-Swahili bilingualism. Consistent with this, qualitative interviews with parents showed that neither Bhele identity nor language were perceived as useful. Table 7 shows language competence of children aged between 6–20 years from Bhele community, in their community's language.

What is immediately striking in this data set is that many children from within the community itself, in its homeland, are not learning even to understand the language, and fewer are learning to speak it. There is a broad parallel with the urban cases we have examined. Furthermore, we note a difference in the rate of loss of children of different ages, and again we cannot, from this data alone, ascertain whether this is due to later acquisition of the communal language, or to reduced rates of transmission over the years. It may well be a combination of these two phenomena. In terms of the EGIDS (Lewis & Simons 2010), the language in these communities would be at Level 6b, *Threatened*, as there are many children who are not learning it.

In our other case studies, in Nairobi and Bunia, we are looking at languages 'imported' into the urban context, and do not conclude that these languages in their home communities are under the same level of threat that Bhele appears to be from this data. We also note that the LWC Swahili has found a place in rural homes, not only in urban ones. We must suppose that this rural shift to Swahili (a language which is not perceived as belonging to one particular ethnic group, and therefore not automatically signifying an identity shift) is facilitated by its widespread use in urban contexts; we have been dealing with urban and rural contexts as different, and they are, and yet they do not exist in isolation from each other, nor is the dichotomy a categorical one (we may need to speak of urban and rural settlements on some sort of a continuum, or a sliding scale).

4.   Final questions

We have noted commonalities between the situations in Nairobi and Bunia – in both cases all children are learning the LWC Swahili, while many, but not all, also maintain competence in the "imported" vernaculars. There is little evidence for a deliberate turning away from the vernaculars, just that they are generally perceived as not having the same value as Swahili, French or English (not that the latter two appear to be used much as home languages). There is little family management to protect the vernaculars, and if there is, then it is not community-wide. What is so interesting is that we have found very similar trends and attitudes in one rural context (Bhele) as well. However, the reason that we report on this context is not its typicality, but its comparative exceptionalism; we do not observe this shift for the majority of languages in their rural "homelands" in Eastern Congo. But the fact that Swahili is so widely spoken in towns makes it an available target for acquisition in the countryside as well. And in this case this move towards Swahili is accompanied by the acquisition of another vernacular, Nande, unlike the urban cases. Is this more evidence that many African contexts are fundamentally multilingual?

There is no doubt that Swahili is gaining many speakers, many of whom will use it as a first language alongside at least one other. We may add that its prevalence in so many urban centres in East Africa raises its perceived benefits (Karan 2011) for speakers across the whole region; a virtuous circle where being useful further increases perceived utility, especially where urbanisation proceeds apace.

The move of Swahili into the home domain in the countryside raises a larger question about language shift and maintenance in Eastern (and perhaps other parts of) Africa. Are we seeing a change from a pattern where language shift in this part of the world is from one communal language to another (e.g. Bagamba 2007, which looks at the shift from Oruhema to Bhadha) to a phase where the shift is primarily in the direction of LWCs such as Swahili or Lingala, which do not have such a strongly associated ethnic identity? This would seem to be a broader trend in other parts of Africa, not just the East. As part of this issue, found in a background of widespread urbanisation, we must ask the broader question of how the dynamics of urban language ecologies in Africa impact those in the countryside. We do not yet know the answer to this question, but this paper is written in the hope that it might stimulate thought and further research which will inform a possible answer in the future.

References

Bagamba, B. Araali. 2007. A Study of Language Shift in Rural Africa: The Hema of the North-East of the Democratic Republic of Congo. PhD dissertation, Essex University.

Bagamba, B. Araali. 2012. Study of Language Shift among the Bhele. *Ms*.

Bagamba, B. Araali & Gibson, M.L. 2014. Urban Diasporas and Intergenerational Transmission of Vernaculars in East-Africa: The Case of Northern Hema Community in Bunia, DRC. Ms.

Bagamba, B. Araali & Gibson, M.L. In press. *Urban diasporas and intergenerational transmission of vernaculars in east-Africa: The Kisii in Homa Bay*; Kenya.

Batibo, H. 1992. The fate of ethnic languages in Tanzania. In Brenzinger (ed.). 85–98.

Brenzinger, M. (ed.). 1992. *Language Death: Factual and Theoretical Explorations with Special Reference to East Africa* [Contributions to the Sociology of Language 64]. Berlin: Mouton de Gruyter. doi:10.1515/9783110870602

Fishman, J.A. 1971. The sociology of language: An interdisciplinary social science approach to language in society. *Advances in the Sociology of Language* 1: 215–404.

Gibson, M.L. 2012. Language shift in Nairobi. In *Proceedings of the 6th World Congress of African Linguistics, Cologne, Germany, 17–21 August 2009*, M. Brenzinger & A.-M. Fehn (eds), 569–575. Cologne: Rüdiger Köppe.

Karan, M.E. 2011. Understanding and forecasting ethnolinguistic vitality. *Journal of Multilingual and Multicultural Development* 32(2):137–149. doi:10.1080/01434632.2010.541916

Kaputo, S. 1982. *Phénomène d'Ethnicité et Conflit Ethnopolitique en Afrique Noire Post-coloniale*. Kinshasa: Presse Universitaire du Zaïre.

Legère, K. 1992. Language shift in Tanzania. In Brenzinger (ed.), 99–115.

Lewis, M.P. & Simons, G. 2010. Assessing endangerment: Expanding Fishman's GIDS. *Revue Roumaine de Linguistique* 55(2): 103–120.

Lüpke, F. & Storch, A. 2013. *Repertoires and Choices in African Languages*. Berlin: De Gruyter. doi:10.1515/9781614511946

Mufwene, S.S. 2010. The role of mother-tongue schooling in eradicating poverty: A response to Language and Poverty. *Language* 86(4): 910–932.

Obedi Poun'ga, W. 2009. Aperçu panoramique de Bunia. Revue Shalom de Développement 1111-2007-32 du troisième trimestre 2007.

Ucoun, C. 2012. Transmission des langues maternelles en milieu urbain: Cas de Dhu-Alur à Bunia. PhD dissertation, ISP-Bunia.

Redefining priorities, methods and standards in endangered-language lexicography

From lexical erosion in Palikur to areal lexicography

François Nemo and Antonia Cristinoi
University of Orléans

Based on our ten-year fieldwork on Palikur (Arawakan) in French Guyana, this paper discusses the need for researchers (assisted by native speakers) to devote more time and energy to documenting the lexicons of endangered languages (EL) and the urgency to focus on the parts of the lexicon that are subject to strong lexical erosion, in order to allow the transmission of saved knowledge to younger generations. It also describes the main difficulties in documenting these sections of the lexicon and advocates the need for new principles, methods and standards in EL lexicography so as to obtain more complete and satisfactory information for interdisciplinary work, for tool, knowledge and content sharing, and finally for dismissing an all-or-nothing approach according to which incomplete lexicographical descriptions and knowledge should not be published.

Keywords: Amazonia, biolexicon, dictionary, documentation, endangered language/s, lexical erosion, lexicography, lexicon, Palikur, transmission

1. Introduction

Based on our experience of language documentation and field lexicography in an Amazonian context, namely our work on a Palikur-French dictionary in French Guyana, our aim here is to consider what can be done and changed in order to avoid *lexical erosion*.

After defining lexical erosion in general and in the Palikur context in particular, and after stressing the importance of preserving the lexicon and its use in specific situations in a language conservation situation, we shall discuss what the priorities, methods and specificities of EL lexicography should be in such conditions and, in order to meet these particular challenges, we shall advocate

a necessity to reconsider lexicographical work and formats.[1] We shall build our reasoning on the following premises:

1. that more time and energy should be devoted to documenting the lexicons of ELs, which is too often considered as a minor scientific objective, and the existing dictionaries are not really formatted to truly contribute to word and word meaning preservation;
2. that we need to adopt an *emergency approach* to field lexicography and language documentation in order to save the frailest sections of the lexicon (that are the most likely to disappear);
3. that the lexical treatment of each lexicographical entry must be reconsidered in order to allow its preservation;
4. that access to the saved lexicon (and knowledge) must be provided for younger generations.

Focusing on biological and environmental knowledge and the *biolexicon*[2] as an illustration of a part of the lexicon that has to be saved before it is too late, we shall then discuss:

1. the specific difficulties in documenting these sections of the lexicon for the average linguist and/or lexicographer;
2. the need to obtain much more precise descriptions of the lexical units involved than is usually considered necessary;
3. the different techniques which may and must be used to obtain these types of information;
4. the necessity to associate the classical lexicographical approach with other methods of lexical documentation, namely collecting small oral corpora for each lexical item;
5. the need for the linguist to engage in multi-disciplinary work when dealing with sections of the lexicon like the biolexicon (plant and animal names to be

---

1. For a presentation of existing formats and standards in lexicographical work, see for instance, Hartmann (2003) and Mosel (2011).

2. As far as Palikur is concerned, the risk for the dictionary to contribute to bio-piracy is almost null, mainly because their use of medicinal plants, for instance, has already been documented in great detail (*Pharmacopées traditionnelles en Guyane*, Grenand et al. 2006). As for the general issue of knowing whether all the collected information should be made public or not, it is important to distinguish between the importance of storing all information and the fact of making it public. Whereas all our Palikur oral corpora, for instance, will be deposited and stored at the BnF (French National Library), researchers and their informants have the possibility to decide on the availability of each corpus.

exact), which cannot be collected thoroughly enough without a certain level of expertise;

6. given the scarcity of resources available and the time-consuming nature of lexical documentation in specific linguistic areas such as Amazonia, there is an urgency for linguists to share their tools and knowledge or to build these tools together;

7. the fact that since it is possible to considerably reduce the time-consuming nature of lexicographical work and to enhance the quality of basic encyclopedic glosses (like glosses describing plants, for instance) if glosses could be shared by all lexicographers within a linguistic area such as Amazonia, it would also be extremely useful to adopt new authorship standards which would allow such sharing practices.

Finally, we shall dwell on the necessity to reconsider the formatting of both micro-structures and macro-structures in EL lexicography, either by meeting the specific lexical description constraints associated with preservation issues, or on the contrary by rejecting an all-or-nothing approach to EL documentation which may lead to the loss of precious but incomplete and thus unpublishable information.

We shall further discuss using phonetic forms (written in IPA) as headwords whenever the existing word spelling may prove ambiguous or whenever it may ensure easier access to the information provided by the dictionary for the members of a given linguistic community.[3]

2.   On lexical erosion

A language may disappear as a whole, but it can also disappear by subject specific fractions, through a continuous process of lexical erosion (see also Cristinoi & Nemo 2013), which takes place whenever important changes affecting one linguistic group rapidly modify the lifestyle of its members (schooling, relation to the environment, medicine, religion and language contact), thus threatening the transmission of both words and word meanings.

In the Amazonian context, the words which are the most likely to disappear are the ones that name biological entities (animals, plants) and environmental knowledge, ritual practices, mythical entities, traditional medication and objects which are no longer used in everyday life.

---

3.   In our case, as the Palikur phonetic system is rather simple, the IPA transcription would be easily readable by a Palikur speaker.

2.1   The Palikur context

Even though the main goal of this paper is not to describe in detail the dynamics of an Amazonian language such as Palikur, some of its features are worth mentioning. Palikur is an Arawakan language spoken in the north of the Brazilian state of Amapa and in French Guyana, representing today a community of possibly 3000 people, most of whom still speak the language (but have a rather uneven linguistic competence). The Palikur community has a rapidly growing demography (at its smallest, it was probably reduced to less than 400 people), a tendency to spread westwards in French Guyana with new and fast growing settlements west of Cayenne (Macouria) and a high degree of population mobility between the different areas, notably from Brazil to French Guyana (for instance, in Saint-Georges de l'Oyapock, possibly half of the Palikur community has come from Brazil during the past forty years and paradoxically Palikur is to some degree becoming a kind of vehicular language between French-speaking and Brazilian-speaking Palikur people). All of this has direct consequences for language dynamics (as a result of marriages between Brazilian and Guyanese Palikur, villages such as Favard, where Palikur has been replaced by Guyanese Creole during the past fifty years, have recently experienced the reappearance of Palikur speakers, with young children speaking the language while some of their grandparents do not).

Palikur is exposed to multilingualism and has a long-term relationship with an old variety of French-based Creole (which is the official language of its Karipuna neighbors and was spoken in the area before Northern Amapa was assigned to Brazil at the start of the twentieth century), and experiences both resistance to creolization and attraction to it. It has faced a more recent but massive exposure to French and Brazilian Portuguese through media and the school system (Cristinoi & Nemo 2014 unpublished field data). The primary school system in the state of Amapa is a bilingual one (Palikur-Portuguese), which implies among other things that writing in Palikur is part of the curriculum, whereas in French Guyana Palikur and other indigenous languages have long been ignored within the education system. French has been the only language of schooling and education for a very long time and vernacular languages have only recently started to be studied in schools on a more or less regular basis.

The dynamics of the language appear to vary in the different areas, depending on the chronology of each settlement. In the oldest ones, notably in Saint-Georges de l'Oyapock, one may notice a recent tendency to adopt a language mix of two main and equally important languages, Palikur and French. In others, a language mix of Creole and Palikur is now challenged by the rising importance of French and the disappearance of Creole-speaking neighbors, such as in Régina.

2.2   Lexical erosion in the Palikur context

Generally speaking, the use and the vitality of a language may be measured in terms of the number of active users (speakers), especially in younger generations,[4] and hence it could be said that the Palikur situation is not dramatic in this respect. Nevertheless, the vitality of a language also depends on the scope of its use and its ability to be used in all areas of life. Much has been said about the capacity of languages, even strong languages, to deal with the necessity to adapt lexically to new patterns of life and to modernity and external contact. As far as the lexicon is concerned, adapting to modernity appears to be one aspect of the problem, since it is impossible to assume that the lexicon of a language is automatically transmitted to the next generations. However, it is also crucial that this happens so that the use of a language may be pursued in various domains. When non-transmission occurs, the generation gap in terms of lexical competence is a reality with which lexicographers are directly confronted. Measuring this lexical competence, however, is a larger issue which requires extensive work and specific criteria in order to be precisely defined, measured, characterized and understood in its variability.

In the Palikur context, after considering all the members of the community, it soon became obvious that lexical competence was uneven, that it was necessary to measure its variability, and that it probably included lexical erosion between generations. As a first step, we have thus developed lexical surveys aimed at measuring actual lexical competence and knowledge among the Palikur with the following methodology (Cristinoi & Nemo 2014 unpublished field data):

1. the surveys concerned a range of distinct lexical fields so as to cover all aspects of life (and the lexicon);
2. for each field, the knowledge of a set of supposedly basic, non-basic and sophisticated (i.e. quite specialized) words was tested;
3. all tests were conducted in Palikur and were bidirectional, either providing the words and asking about their meaning or by providing notions and asking for their Palikur equivalents;
4. the survey was conducted in all the major Palikur settlements in French Guyana (notably Saint-Georges de l'Oyapock, Régina and two villages in the Macouria area: the Kamuyene and Norino villages);

---

4.   It must be noted that the reluctance of teenagers to speak the language should not be considered as definitive. In Saint-Georges de l'Oyapock, for instance, the teenagers from the last years of the twentieth century were often simply refusing to admit speaking Palikur but many of them have later become language-transmitting parents.

5. the survey was conducted for different age and sex categories in each settle-
    ment and the personal profile of each interviewee was documented, in terms
    of movements between villages and from Brazil to Guyana and also in terms
    of education and schooling.

These surveys have shown that the differences/discrepancies and commonal-
ities between speakers and settlements follow different patterns. To start with,
we noted that lexical erosion and the loss of vocabulary is not always the result
of modern acculturation or language disappearance but is often the quick result
of the existence or inexistence of specific techniques and practices in the vari-
ous settlements of the community, since the learning of specific words appears
to be greatly conditioned by the actual use of the corresponding objects. If such
objects are sometimes simply not relevant in the specific environment in which
a new settlement appears, the words may disappear from lexical consciousness
very quickly, as could be observed in Saint-Georges de l'Oyapock for large parts
of the fishing vocabulary, due to the impossibility of using the fishing techniques
found in Rocawa (Amapa) and in other settlements; this is why it is crucial to
document them.

  Another crucial observation is that the patterns of lexical erosion which could
be observed do not follow a pattern of *reduction* to a core (or fundamental) Pa-
likur lexicon, in which all *basic words* would have remained known, followed by
less-remembered *non-basic words* and finally a forgotten *specialized vocabulary*. It
appears on the contrary that reduction is massively field related: for instance, in
the villages in which the overall lexical erosion was the strongest, we could find
young Palikur who seem to know quite well the lexicon of family life, including
idiomatic expressions, but are almost unable to name any animal in Palikur, in-
cluding those living on their doorstep like cats.[5]

  It also appears that lexical erosion is a highly generational process, which
means that at certain moments, important portions of the lexicon still exist
among the older generations (50+) but could be lost for the younger ones, whose
life has deeply changed, and this is independent of their global attitude toward
the language itself.

  As for Palikur, it appears that despite the fact that the overall number of Pali-
kur speakers has probably never been higher during the past centuries, there is a
big risk and strong probability that in the next twenty-five years half of the Pali-
kur lexicon will be gone, and, as importantly, that the impoverishment of certain

---

5. It must be noted that such lexical surveys complement ethnographic work as they allow
new access to the reality of language use within the diverse settings of a community. In our case,
years of fieldwork has revealed surprisingly diverging lexical dynamics among these settings.

lexical fields could simply block the use of Palikur for evoking aspects of life which paradoxically have always been central to the Palikur culture.[6]

2.3   Dictionaries and lexical erosion

Without considering this issue in full detail here, it seems that such patterns of lexical erosion and lexicon endangerment have dramatic consequences for language endangerment as a whole. The implication is that beyond a language's capacity to provide vocabulary for the modern world, its capacity to resist such lexical shift/lexical replacement is crucial for its future.

In this respect, what appears crucial whenever lexical erosion has already started is the capacity to restore and revitalize a language's lexicon and in order to do so the tools which allow an indirect transmission of the lexicon need to be created.

It must be stressed that for a language such as Palikur, the generational gap associated with literacy and spending years in the school system, even though it endangers the traditional oral transmission of the lexicon, also enables dictionaries and websites to become a new form of collective memorization of words and word uses. This must stimulate researchers to find new modes to transmit knowledge and experience between generations.

However, for a paper or online dictionary to play this role, much has to be reconsidered in the elaboration of the dictionary itself, as we shall see further on in this paper.

3.   Reconsidering methods in field lexicography and EL lexicography

Field lexicography for minority languages or endangered ones is in itself a difficult task, which faces many limitations. If it were to be considered in terms of preservation of the lexicon, its principles would have to be reviewed.

3.1   Limitations of EL lexicography

Some of the shortcomings of minority or EL lexicography are global, and among them we can note:

---

6.   A complete presentation of the Palikur people and culture is available in Grenand (2009). For a presentation of the grammar of Palikur, see Launey (2003). The only published Palikur dictionary is Green & Green (2010).

1. the fact that it is a time-consuming task;
2. the fact it is under-funded;
3. the frequent absence of available data and corpora, and its numerous conse-
   quences, among which we find the under-representation of lexical polysemy;
4. the shortcomings of the various existing techniques of lexical elicitation;
5. its reliance on mostly individual work;
6. the consequent lack of expertise of the lexicographer in many fields of knowl-
   edge, and specifically in environment, culture or area-specific aspects of life.

Nevertheless, some of the shortcomings are purely lexicographical (even when
they are ultimately related to global ones), such as:

a. the fact that the description of words for many entries may be considered as
   insufficient or suboptimal;
b. the fact that the phonetic form is often absent in existing paper dictionaries
   and frequently unpredictable from the orthographic form and that few re-
   cordings and electronic dictionaries (other than mere electronic versions of
   the paper dictionaries) are available for these languages;
c. the fact that the coverage of certain lexical domains is often minimal;
d. the fact that aspects of a lexical entry which could be considered as especially
   relevant for a given language are not included.[7]

The direct consequence of such shortcomings is, as mentioned earlier, that field
lexicography in its current state is often unable to really fulfill the goal of preserv-
ing a lexicon and fighting lexical erosion.

To a large extent though, this limited capacity to avoid lexical loss and to stop
lexical erosion is also due to the fact that what lexicography ought to achieve in
an EL context has not been clearly spelled out and transformed into a genuine
constraint on lexicographical work, neither in terms of methods, nor priorities
or content.

This is why our aim here will be to try to formulate areas in which changes in
goals, methods and sometimes structures could significantly improve the capac-
ity of field lexicography to truly contribute to language preservation and lexicon
transmission.

---

7. Such as the specific numeral classifier used for a given noun in Palikur, even though it is
especially important to ensure the effective use of such a word in context. This concerns a quite
unique feature of Palikur in terms of linguistic typology and appears to be suffering strong
generational loss; documenting it item by item should be considered compulsory.

3.2   Principles and priorities in EL lexicography

Bearing in mind the fact that producing a dictionary is a long and undervalued task, the result of which is as easy to criticize as it is difficult to achieve, and that our remarks or illustrations of the issues at stake should not be considered as a form of bashing the considerable work that they represent, we shall now try to consider what a dictionary should be if we want it to play an effective and significant role in the preservation of the lexicon.

We can state some basic principles concerning the priorities of EL lexicography, both at the level of individual languages and individual lexicographers, and at the level of countries or linguistic areas.

The first principle that must be implemented in order to fight lexical erosion is, as mentioned above, to consider the frailest part of the lexicon as a priority in lexicographical work. This may be called *the emergency principle*.

The second principle is that all should be done as if the dictionary was to become the only available source of information on the words or on the language at stake. This may be called *the exhaustivity principle*.

The third principle is that whatever can be done to reduce the time required to meet the requisites which follow from the first two principles should be done. This may be called *the economy principle*.

The fourth principle is that whatever can be done to enhance the quality of what is produced should be done, even if it deviates from the way lexicographical work on non-endangered languages is usually done. This may be metaphorically called *the ISO principle*. The idea behind this "ISO" principle (the norm referring here to classical ISO standards) is that we should have specific quality standards for this type of lexicography, standards that would be different from those of classical lexicography.

The last principle is that whatever can be done to ensure transmission – and not only scientific documentation per se – to the next generations should be done. This may be called *the transmission principle*.

Adopting and reconciling all these principles as both specific to EL lexicography and central to it has huge consequences for the methods, standards and formatting of the lexicographical work and the dictionaries themselves.

3.3   Methods in EL lexicography

The first consequence of the priorities and principles mentioned above is to concretely ask much more from the lexicographers of an EL than is required for a non-endangered language – in terms of exhaustivity of the lexicographical coverage of a language, the inclusion of portions of the lexicon whose treatment is

very demanding, and in terms of the precision of the information provided on each item.

The second consequence is that because raising the standards that have to be met when writing a dictionary to ensure actual preservation of the lexicon is a necessity that cannot be satisfied by individual lexicographers most of the time (with an individual lexicographer being either a trained researcher, linguist, ethnographer or a missionary or a trained native speaker), it is therefore necessary to considerably help them in doing so, but also to move EL lexicography away from some of the limitations of authorship. It is necessary to adopt a much more collective and collaborative framework based on multi-disciplinary, areal approaches and on lexicographical equivalents of freeware.

3.3.1   *Facing the lexicographer's limits*
The most demanding aspects of the work (from elicitation tools to actual lexicographical descriptions) can indeed be transformed by collective and collaborative work. If we consider, for instance, the biolexicon, based on our experience and the study of dozens of existing dictionaries of the same linguistic family (Maipurean/ Arawakan) or area (Amazonia), the reality is that existing standards or practices are well below what would be necessary to allow preservation, for a number of reasons which are easy to understand, notably:

1. the lexicographer's own limits when it comes to dealing with realities he/she has no knowledge of;
2. the time needed to acquire such knowledge;
3. problems in elicitation techniques (using pre-determinate word lists and translations, and single stimuli like drawings or pictures, for instance, which lead to numerous errors especially as far as small animals and plants are concerned).

All of this leads to the fact that most of the time the authors of EL dictionaries cannot reasonably provide what is asked of them.

3.3.2   *Sharing resources and lexicographical glosses*
It follows from this that what has to be done cannot be done autonomously for each individual language, due to a lack of financial and human resources, and thus the only way for EL lexicography to meet its goals and to limit lexical erosion is to deeply transform the way field lexicography is carried out and conceived, and specifically to stop asking each individual lexicographer to deal with specialized domains such as the biolexicon in which he/she has little competence.

Fortunately, regarding this last point, it is clear that for large portions of the lexicon what has to be done could be done for all the languages of a given area or country if resources were fully shared.

Even though it is solitary work, EL lexicography would gain considerably if researchers made available, for each linguistic area, a kit of tools and knowledge which could be used by each lexicographer to work more efficiently and therefore limit the amount of time required to obtain satisfactory results in the description of the biolexicon. For example, the creation of a (animal) sound database for Amazonian lexicography would allow for work on dozens of languages to be far more efficient. Similarly, given the areal nature of many names, a collaborative compilation of existing knowledge would considerably simplify etymological issues.

3.3.3   *Sharing resources: Expertise*
To stick to our study of lexicography of the biolexicon among the Palikur, even though we had a non-basic knowledge of some parts of it, collaboration with biologists, ethno-linguists, ethno-pharmacologists and anthropologists indeed proved to a necessity, allowing both sides to obtain:

a.   much more data;
b.   much more precise data (in both biological and orthographic/linguistic dimensions of the issue), in a much shorter period of time.

An example of collaborative work of this kind is a two-week field trip conducted in 2011 with archeologists, botanists, anthropologists and Palikur informants, which consisted of:

–   the identification and tagging (with number and scientific name) by botanists of 800 trees of an archeological site in the forest;
–   the identification and naming by Palikur informants of each one of those 800 trees;
–   the recording of short (2–3 minutes) ethno-linguistic comments (in Palikur) on each tree-type and its uses (including stories and information on traditional architecture, boat building and traditional object making according to the specific uses of the trees), thus allowing us to create a large corpus which could eventually be made available in online versions of the dictionary, and thus made available to the next generation;
–   the use of this corpus to extract other unknown lexical items or unknown meanings of known ones.

Another joint research project was conducted on insects and ethno-pharmacology, leading to results which are crucial for the illustration of the dictionary but which could also be used for areal purposes.[8]

3.3.4 *Sharing resources: Elicitation tools*

When working on the description of lexical fields which are not familiar to him/her, the lexicographer cannot but rely on accessible tools (pictures, drawings, sounds, videos, etc.) and is nevertheless confronted with the difficulty of producing reliable information.

As far as the biolexicon is concerned, any linguist who has tried to obtain reliable denominations for specific animals, birds and plants knows how important and troublesome elicitation techniques are, even though the multiplication of field guides or websites about almost any kind of animal and plant are making it easier nowadays than it used to be.

Anyone wanting to use sound files to elicit bird names, for example, can use a website such as *http://www.xeno-canto.org* where he/she may easily obtain geographically relevant files of bird songs to be tested. This allows one to obtain names which otherwise would remain unknown or would be wrongly attributed.

Using drawings or pictures to elicit names appears to be a technique which, despite its obvious advantages for the lexicographer in terms of exhaustivity and precision, faces strong limitations, among which is the fact that form is often not the main criterion of animal identification in the real life experience of the speakers, the fact that size is virtually impossible to represent through images, and more importantly, the fact that for huge parts of the biolexicon (insects, trees, snakes to a large degree, frogs to a significant degree, etc.), this technique appears impracticable.

This is also tricky because of the fact that if A may be called N, this should never mean that N actually applies only to A, or the fact that a Western hyperonym H (general term indicating the class of objects the noun belongs to) is called N by an informant should never mean that N has indeed an hyperonymic status in his/her language, and also and most importantly, because presenting in a L2/L1 order what has been obtained by a L1/L2 question can be misleading. Indeed, linguistic inquiry based on images is always associated with verbal explanations from the inquirer which can deeply affect the nature of the answers obtained and can, for instance, produce artifacts and/or mere translations. Similarly, among

---

8. Because bees, for instance, are to a large extent named in Palikur according to the type of nest which they build, having a picture (or a verbal description) of the different kinds of nests is important for the dictionary but could also serve as a basis for the definition of an areal bee Swadesh-like list.

polyglot informants, the interference between naming systems may be so strong that extreme caution is required.

This is why direct field experience may prove to be crucial: after two complete "non-field" studies of Palikur birds by two distinct researchers (including sound use in one of them), a simple fifteen-minute walk provided no less than four unknown names of familiar birds, a one-hour trip to distinct biotopes can provide dozens more, with the only, but sometimes demanding, necessity being to identify the birds at stake.

While all this does not make things easy for the lexicographer, it is nevertheless true that simply knowing where to find the tools for the task allows one to considerably reduce the time needed to document a group of animals or plants and to increase the number of names that will be included in the dictionary.

But once again, it is obvious that if all the necessary preparation of the task was not left to the individual lexicographer but mutualized at country, or preferably area, level, it would become easier to document parts of the lexicon which until now have appeared too difficult to document. Having at one's disposal on a website all the shapes of bee/wasp nests in Amazonia, for instance, would allow extensive coverage and precision for dozens, if not hundreds, of minority languages. This proves, once again, that efficiency at the level of an individual language would be greatly enhanced by sharing resources at area level, at no extra cost.

3.3.5   *Sharing resources: Lexicographical glosses*

When the lexicographer of an individual language has managed to obtain all the data needed to fully include the biolexicon in his/her dictionary, he/she is left with the purely lexicographical task of describing each lexical item, in other words, producing specific lexicographical glosses and definitions for each word.

Without going into further detail, it appears from observing the way this task is performed in existing dictionaries that it is not an easy one for two main reasons:

– good definitions are not only true statements but should also provide sufficient information (see next section);
– good definitions need to be adapted to the user/s of the dictionary.

Regarding this last point, a difficulty of the task is that dictionaries of minority or endangered languages are most of the time bilingual dictionaries, with potentially three types of users: speakers of the majority language (school teachers, doctors or ordinary citizens), speakers of the minority language itself, and finally, scientists (Cristinoi & Nemo 2013). The definitions and glosses found in existing dictionaries thus oscillate (sometimes within a single dictionary) between the

three perspectives, often adopting only one of them, which makes them opaque or irrelevant for the other users.

Our claim here is that the main objective of a dictionary should be language (and culture) preservation and in order to secure it one should combine in each individual word definition/description all the information which would ensure lexical transmission and avoid lexical fuzziness.

As for the biolexicon, the following is an example of what may happen for each individual lexical item in a contemporary dictionary. It is well known that for certain Australian languages, only a small list of nouns could be documented before the language stopped being used. For instance, in a case where a name for "heron" was documented, it is actually the case that given the existence of no less than seventeen species of herons in the area, the contemporary descendants of the community who are making language revitalization efforts will never know if the term applied to a specific heron species, or whether it was a hyperonym. Quite too often, a similar pattern may be found in existing dictionaries: what the definition/gloss provides is half of the information needed.

But since providing optimal glosses or definitions is a difficult and time-consuming task, it is natural that most dictionaries have to adopt lower standards, and in a situation of lexical erosion face the loss of non-provided information.

As for the biolexicon, the issue appears once again unsolvable at the level of an individual language, and not only because of the workload, but because of the fact that if "good" definitions/descriptions/glosses of let's say all the Amazonian birds (of a given area) were ever created, they would remain (as far as publication is concerned) the property of the lexicographer who has provided them and would not be usable by other lexicographers/dictionary compilers lest they should be accused of plagiarism.

This is a strange paradox, since in many countries dozens of languages face the same problem, with little means to deal with it. They could benefit considerably from the availability of good lexicographical descriptions of animals or plants but cannot do so because of a conception of lexicographical work centered on single authorship and the risk of copyright-based accusations.

In a continent like South America, enhancing the capacity of dictionaries to reach a level of description of lexical items which allows the effective and full preservation of the words at stake would greatly benefit from making available ready-to-use lexicographical glosses[9] for sections of the lexicon such as the biolexicon.

---

9. For the Amazonian region, such glosses should become available in Spanish, Portuguese, French, English and Dutch. Because the way the gloss is written corresponds to a communicative strategy, its pattern allows the use of translation to make the glosses available in the five languages.

This does not mean that the individual lexicographer may not make up his/her own glosses or adapt or improve these freely available glosses, but only that given the scarcity of resources available for the creation of dictionaries of so-called small languages, sharing such lexicographical resources could be the only way to produce them.

3.4   Reconsidering standards and structures in EL lexicography

EL or minority language lexicography is too often considered in terms of negative comparisons with Western lexicographical standards that have emerged as the result of centuries of constant and renewed work and with the financial backing of powerful publishers. Such a perspective of course is partially grounded in huge and undisputable contrasts, in terms of the possibility to dedicate financial or human resources for lexicographical work, but has its shortcomings.

The main shortcoming of approaching small language lexicography as a substandard lexicography whose ultimate goal would be to resemble as much as possible the lexicography of the main languages of the world is that it does not question the assumption that small language lexicography may require a standard of its own, adapted to the sociolinguistic and cultural situation of the language and to its users and aimed at specific goals, such as preservation, which are mostly irrelevant in the lexicography of major languages.

3.4.1   *Beyond orthographic nomenclatures*
A very simple illustration of this subject in the Palikur situation is the issue of deciding whether the addressing system of the dictionary should or should not be orthographic in a context in which nearly half of the Palikur live in French Guyana where:

– the writing and spelling of Palikur is not taught in the school system (whereas the other half of the population is educated in Brazil in a bilingual primary school system);
– the existing orthographic standard, due to various circumstances which will not be detailed here, is to some degree opaque and unpredictable;
– many of the words that will be included in the dictionary have never been written so they have no standard spelling.

It thus becomes impossible to *presuppose* knowledge of word spelling and to rely on it for the ordinary user to be able to simply find any word in the dictionary. This leaves the lexicographer with the choice of sticking to standard alphabetical ordering of spelled words (used in Brazil) with the consequence that most

potential Palikur users will be unable to use the dictionary as it is[10] or will have to rely on the French-Palikur section to find the right word spelling (supposing they already know the spelling of the French word), or with the choice of permuting phonetic transcription and word spelling (using the phonetic and not the orthographic form as an entry in the dictionary), thus allowing access to the latter through a simplified version of the former, which can be easily guessed and can thus ensure the usability of the dictionary.

Leaving aside this illustration, it is obvious that once one starts asking oneself what could be changed in classical lexicographical standards in order to allow the dictionary to provide as much relevant information as possible – and moreover if this question is raised according to the various principles we have defined to allow language transmission and preservation – it becomes clear that it will only take a few seconds for the lexicographer to record his/her informant's pronunciation of a given word, and it will take him/her only a few minutes to record his/her informant speaking in Palikur (and possibly also Creole/French) about a given bird or plant. This will create thousands of small corpora which can be archived and thus remain accessible for the next generations and allow not only the documentation of a word but of the cultural representation and experience associated with it.

3.4.2 *Written examples*

Going back to the issue of whether or not EL lexicography should have its own standards, it is obvious that recording the pronunciation of a word and the knowledge associated with it (including the use of each word in context) may be far more important than creating examples for each word (the practice of standard dictionaries) in a context of lexical erosion, even if the latter is somewhat useful. Ultimately, having small recordings associated with thousands of individual words (or meanings) is not difficult to achieve – given that the lexicographer is documenting each word anyway with his/her informants – and may prove to be a new and efficient way to ensure the transmission of the knowledge at stake.[11]

3.4.3 *General constraints on lexicographical content*

Discussing standards, however, goes well beyond discussing the exact structuring of an entry, as lexicographical content and work are ultimately ruled by the same

---

10. Orthographic bidirectional tests have been conducted in order to measure this capacity, showing that it was actually lower than what could have been expected.

11. This of course implies moving away from the paper dictionary perspective, but once again, with mobile phones spreading in the Palikur community as in anywhere else in the world, transmission could well take the electronic path.

pragmatic constraints that apply to any contribution in language, namely the Gricean constraints of exhaustivity, clarity, truthfulness and relevance (Grice 1975).

Applying Grice's maxims (and Ducrot's law of exhaustivity) to defining what a lexicographical entry *has to say about something* requires understanding a dictionary or a dictionary entry as a contributional object which has to respect the following maxims:

1. *quantity*: "do not provide less information than required, do not provide more information than required";
2. *quality*: "do not say things for which you haven't got enough evidence";
3. *manner*: "do not say anything obscure";
4. *relation*: " be relevant".

It is indeed these general constraints that make the lack of exhaustivity of a dictionary somewhat unsatisfactory in terms of coverage of a lexical domain, as is the lack of exhaustivity of a lexicographical entry in terms of the information it provides in contrast with the information that is (or could be) expected – this second issue having strong consequences on the first one.

An important challenge in this respect and in the preservation perspective that we have adopted here is to combine such general constraints with the principles that we have introduced earlier. In other words, in the context of EL lexicography, defining the actual meaning of *do not provide less information than required* equals defining clearly what information is required, and defining the actual meaning of *do not say anything obscure* involves defining obscurity in relation to the different types of users, and dealing with the difficulties of satisfying various constraints simultaneously. This can be illustrated, for instance, by the situation where the name of an animal is described only by providing its scientific Latin name, this information being required for precise identification, even though it will be obscure for most users, because it is not common ground. In contrast, providing a common-ground description (such as *heron*) allows the maxim of clarity to be respected but not the maxim of quantity, as there is a lack of precision concerning which kind of heron we are referring to.

Reconsidering the standards in EL lexicography thus means clarifying what information should be provided, or to put it the other way round, clarifying which information is missing or is inadequate, if we want dictionaries to fully serve as tools for lexicon preservation and transmission.

3.4.4 *Specific constraints on biolexicon lexicography*

This can be illustrated by genuine examples of partial satisfaction of the constraint of completeness in some lexical entries of various dictionaries, all of which may be considered as providing good or rather good information about the biolexicon.

If we consider the dictionary of Yukuna (Arawakan):

> *Yukuna Dictionary* (Schauer et al. 2005)

(1) a. *aroʾojómaji, ajoʾojómaji* s. piconcito: picón.[12] V.apén. aves.[13]

                                                          (Schauer et al. 2005: 31)

    b. *iñapimí* s. piconcito: picón. V.apén. aves.   (Schauer et al. 2005: 46)

    c. *choro* s. golondrina.[14] V.apén. aves.   (Schauer et al. 2005: 35)

    d. *juripíchaʾa* s. golondrina. V.apén. aves.   (Schauer et al. 2005: 67)

    e. *jutá, utá* s. golondrina negra.[15] V.apén. aves.   (Schauer et al. 2005: 67)

    f. *juwiche* s. azulejo[16] (especie de tángara). V.apén. aves.

                                                            (Schauer et al. 2005: 67)

    g. *kaijmeru* s. mariposa.[17]   (Schauer et al. 2005: 70)

       kaijméruna (término genérico) mariposas.

       lachamaru mariposa nocturna.[18]

       paʾajrú mariposa nocturna.

       phichí panami mariposa grande, color café.[19]

       pina mariposa grande, morfa azul.[20] V.apén. insectos.

    h. *kalapichi perí* s. gavilán especie.[21]V. perí. V.apén. aves.

                                                            (Schauer et al. 2005: 73)

    i. *kapana* s. gavilán, especie más grande.[22] V. perí. V.apén. aves.

                                                            (Schauer et al. 2005: 77)

---

12. Small woodpecker, woodpecker. The standard Spanish for woodpecker is nevertheless *pájaro carpintero*.

13. Bird.

14. Swallow.

15. Black swallow.

16. From azul/blue.

17. Butterfly.

18. Nocturnal butterfly.

19. Great butterfly with a coffee colour.

20. Great butterfly, blue morpho.

21. Kite species.

22. Kite, bigger species.

Moving now to a *Yine/Castellano Dictionary*:

(2)   a.   *pushchopu* s. victor díaz (especie de pájaro[23])

(Urquía Sebastián & Urquía Sebastián 2008: 21)

b.   *pushropushro* s. pichihuichi (especie de pájaro)

(Urquía Sebastián & Urquía Sebastián 2008: 21)

c.   *shaniyaka* s. huanayo (especie de garza[24])

(Urquía Sebastián & Urquía Sebastián 2008: 33)

d.   *shawashka* s. paucarcillo (especie de pájaro) *Serolnikta nikanata sha-washka*. El paucarcillo estaba comiendo maduro.[25]

(Urquía Sebastián & Urquía Sebastián 2008: 33)

e.   *taki* s. tibe (especie de gaviota[26]) Ksatu gajerni taki. El tibe vive en la playa.[27]   (Urquía Sebastián & Urquía Sebastián 2008: 24)

we may observe that:

– no literal translation of the noun is provided;
– no Latin name is provided;
– local Spanish names are provided, some of them being loans from other Amerindian languages (e.g. *pichihuichi*);
– class membership is provided but at various levels, ranging from direct hyperonyms (e.g. *gaviota/gull*) to the equivalent of what scientific classification calls class (e.g. *bird*);
– class membership is inserted as an explanation of the locally accepted name, *a victor díaz is a species of bird* or a *paucarcillo is a species of bird*, thus providing a piece of information which cannot be considered as common ground and may be said to provide *a piece of information that is required*, but may nevertheless not provide *all the information that is required*.

As for this last point, it may be the case that the ignorant reader does not know what kind of bird a *victor díaz* or *paucarcillo* actually is, and may need another dictionary to find out, and that every time the concerned species will not be a familiar one, which may be the case for most living beings. Nevertheless, all the

---

23. Bird species.

24. Heron/Ardeide species.

25. The paucarcillo is eating wood.

26. Gull species.

27. The tibe lives on the shore/beach.

information provided remains fully relevant, and could ultimately allow the exact referent to be found.

As a contrast, we may now consider the dictionary of another Arawakan language from Colombia, namely Piapoco.

>*Piapoco Dictionary* (Klimpp 1995: 3)
>
>(3) a. *àapi* s. culebra[28] (en general).
>
>   b. *àapi àleu minali* s. macabrel[29] (lit. culebra arbol morador[30]); *Corallus enydris enydris.*
>
>   c. *àapi íinu wiriichu* s. saltòn del monte (lit. culebra su esposa saltamontes[31]); *Tettigniidae.*
>
>   d. *àapi íiwitami* s. vinegarone (lit. culebra cabeza muerte[32]); *Thelyphonidae.*
>
>   e. *àapi wéetéeriwa iyú* s. culebra venenosa[33] (lit. culebra por medio de lo cual morimos[34]).

Leaving aside the last item,[35] we may notice this time that each entry associates three types of information, namely *denominative equivalent, meaning in the language itself* (Piapoco) and ultimately the *Latin/scientific identification* of the species or family. In some respect, it is more complete and satisfactory than the previous descriptions, but with the risk that whenever the denominative equivalent is not a common ground item (*culebra*) like here, providing only a denominative equivalent will not give sufficient information to the reader, who will know how something is named in three languages, Piapoco, Spanish and Latin, and what it means in Piapoco, but will ultimately not know what it is. This could have been the case in Yine for *taki*, whose denominative equivalent *tibe* is not common ground in Spanish and would have remained obscure if the information *a type of gaviota* had not been immediately provided.

---

28. Snake.

29. Tree boa, macabrel common.

30. Lit. Snake tree morador

31. Lit. Snake its spouse grasshopper.

32. Lit. Snake head dead.

33. Venomous snake.

34. Lit. Snake by which we die.

35. Which may be a usual name or might prove to be only a translation from Spanish if obtained through a L2/L1 question.

Such a structure is the consequence of the fact that when it comes to the biolexicon:

– all nouns are also names (e.g. *bluethroat* or *awiyaybu* in Palikur is a noun and the name of a bird);
– they have a meaning as nouns, independent of their meaning as names (*bluethroat* does not mean *bird*, it means *blue throat; awiyaybu* does not mean *bird*, it means *heavy tail*);
– what they name must be identified (a *bluethroat* may be identified as *Luscinia svecica*, an *awiyaybu* may be identified as *Colonia colonus*);
– what the noun stands for must be clarified, which often means decomposing its morphological structure and always means describing its motivation (*cuckoo* cannot be decomposed but refers onomatopoeically to the sound produced by the bird. Since *victor díaz* in Spanish refers to the bird's song, knowing what *pushchopu* means is a necessity);
– it is a necessity, beyond names and nouns, to describe what something *is* by providing the best possible definition of the denominative referent (a *small robin-like passerine bird* for *bluethroat*, a type of *gaviota*/*gull* living on shorelines for *tibe*).

What is important is to realize not only that all three types of information are required, but that the denominative equivalent is often provided in two forms (here Spanish and Latin).

Hence, three compulsory types of information should always be provided and one should be distinguished from the other in order to avoid any ambiguity of their nature:

– denominative equivalent(s) for the identification of the referent;
– lexical or morphemic decomposition for understanding the expression itself (in Frege's sense) and beyond that for understanding its motivation;
– encyclopedic description/definition for a characterization of the referent (as in classical monolingual dictionaries), providing basic synthetic information about the nature of what is named and which makes it accessible to the reader.

This information should be provided even if it is not always easy:

– to identify what exactly a name refers to when it comes to non-spectacular or familiar species;
– to find denominative equivalents in the main language;
– to avoid mismatching between different classificatory levels (i.e. presenting as the name of a class what is in fact only the name of a member of a class);

- to ensure morphemic or lexical decomposition, whenever word-formation is not transparent;
- to be sure of the motivated character of the word.[36]

Moreover, because the three types of information are relevant, it is necessary to achieve clarity to allow them to be discriminated from one another. Otherwise, as mentioned for the case of *golondrina negra*, the reader will be left having to guess whether *golondrina negra* is a translation, the actual name of the bird in Spanish or a short encyclopedic description.

Since a whole article would be necessary to address the issues related to each of these difficulties, we shall limit ourselves here to highlighting some specific issues.

First and most importantly, because of the almost impossible task for the ordinary lexicographer to solely produce thousands of encyclopedic characterizations/definitions which can combine correctness and accessibility to all readers, it would be a major boost for field lexicography if these definitions could be shared and made available for everyone within a country or area.

Such availability would leave each lexicographer with the possibility:

- to use them or not;
- to improve them;

but also:

- to sometimes avoid, through these definitions, imposing one's worldview on another.

It may, for instance, be the case that all birds classified as *avakni* in Palikur can be described as *rapaces* (raptors) in French, but it is also the case that some birds, such as the *kumak* (swallow-tailed kite, also a raptor), since they are not considered as raptors by the Palikur, should be described using a broader term such as *oiseaux* (birds) and/or that the description should mention the discrepancy as such (*Not considered to be a raptor*).

Similarly, it may often be useful and important to combine classificatory systems in the definition in order to overcome the mismatch between categories.

---

36. The term "motivated" refers to the fact that the name given to an animal, for example, is not arbitrary. For instance, the name "heavy-tail" (*awiyaybu*) is motivated by a physical/ethological feature of the bird. Motivation may be transparent, as in this example, or less so, for instance, when hummingbirds are named by the name of a plant because of the similarity of shape between their nest and the plant at stake.

For instance:

– by using monolingual descriptors, for example, describing in Palikur a *wiki* as "a small *titup*, flycatcher-like bird";
– by acknowledging discrepancies, either indirectly (e.g. *tuumwi*: name of a group of birds which includes all woodpeckers, treecreepers and some treecreeper-like ovenbirds/*Furnaridae*) or directly (e.g. *swiswi*: all waders/ *Scolopacidae* except the *sakaska* giant snipe; *inuiri*: identifiable as harpy eagle juvenile but considered a distinct bird).

As for this last point, it must be stressed that if ultimately the constraint of relevance implies that all definitions be adjusted to the targeted audience (community members, local and non-local outsiders, scientists) – and thus finding compromises – prioritizing transmission of cultural knowledge and representations from generation to generation implies including (and favoring) them as much as possible in the encyclopedic definition provided (e.g. *mamatki*: vocal bird which can be heard only during the night and is widely feared and considered as a bad omen).

Ultimately, however, ensuring the transmission of such knowledge will clearly be more effective if it can be recorded in the first place during the lexicographical fieldwork and made available later on. Saving such knowledge has become quite easy and making it available can be achieved in the Palikur context in any school library or on any cultural association's computer. All conditions are thus met to allow the preservation of such knowledge for future generations.

3.4.5   *Sometimes adopting a non-Gricean perspective*
A last and paradoxical issue at this stage concerns the need in EL lexicography to also adopt a radically non-Gricean approach to endangered languages.

Very often, when working with informants, the lexicographer is exposed to what we can call *the partial knowledge syndrome* (PKS), in other words, to having less than complete information about a word.

For obvious reasons, one is then tempted, because of the missing part, to not mention the information at all, in order to avoid criticism or to look incompetent or ridiculous, with the result that too often, a large amount knowledge is finally lost.

This implies that in EL lexicography there is a strong necessity not only to refuse an all-or-nothing approach to language documentation according to which lexical descriptions should not be published if incomplete, but to create a section dedicated to documenting all such forms of partial knowledge.

It may, for instance, happen that we as lexicographers have only heard the name of a (yet unidentified) night bird but cannot provide its identification nor

precise description, and yet as EL lexicographers, we should not exclude this name from the dictionary, nor the little we know about it: storing it in a partial knowledge section could allow the entry to be improved at a later stage, provide the impetus to look for the missing information and eventually save the word.

Moreover, for obvious reasons, vanishing words and word uses being much more exposed to PKS than other words, the paradox is indeed that in many cases, for such "on the edge" words, the little we know will finally be forgotten and lost due to considerations of professional reputation. In Palikur, this concerns a large part of the *kyaptunka* vocabulary (ceremonial and vehicular variety of the language), especially the meaning of certain words found in songs recorded in the 1970s that appear to be lost.

In some cases, incompleteness may be limited or associated to residual uncertainty, as for example when a bird or bug name is known but its exact identification is not yet available, as was the case, for instance, in our Yukuna examples for various nocturnal butterflies (mariposas) and non-specified *picón* woodpeckers (*aroʼojómaji/iñapimí*).

In such cases, meta-communication would be satisfactory, a comment like "not yet identified" replacing the unavailable Latin or local identification. But providing identification clues when they are available would be better (e.g. *A hummingbird with atypical suspended nest*), for it could lead to later identification.

The same thing could be said regarding uncertainty (and the maxim of quality) or residual uncertainty, which should be mentioned as such. For instance, in Palikur the name *pakapka* should be noted as "most probably Pompadour Cotinga, *Xipholena punicea*", so long as the identification is not fully confirmed.

But in cases where incompleteness is more radical, we suggest that the solution of adopting explicitly an "emergency" approach to EL lexicography be implemented – according to which a specific section of the dictionary would be dedicated to words whose knowledge has remained too limited to allow full integration but whose existence has to be documented in order to make possible later recovery.

In other words, dictionaries should accept presenting state-of-the-art lexical knowledge and not exclude nor mask partial knowledge.

4.   Conclusion

What we have tried to show here is that because lexical erosion is a crucial issue for the survival of languages and an aggravating factor of endangerment and because it is often related to lifestyle changes which threaten word learning and

lexical knowledge, new modes of transmission of the lexicon must sometimes be considered in order to allow its preservation.

Furthermore, we asserted that because dictionaries are the usual form of representation of the lexicon of a language, they ought to play a major role in avoiding or countering lexical erosion, but also that in order to ensure actual preservation and transmission of lexical knowledge, field lexicography and EL lexicography have to set their own standards and reconsider their approach to issues such as lexicographical coverage of all lexical fields, completeness of the lexicographical descriptions proposed, usability and readability of the dictionary by community users, enlargement and oralizing of the former "example" section, phonetic form preservation and hearability, recording of the cultural knowledge associated with each item, avoidance of ethnocentric presentation of the items, and creation of a specific section for the documentation of partial lexical knowledge.

What we have finally advocated is a conception of lexicographical work in which the individual lexicographer is not asked to do everything. Rather, instead of hundreds of isolated and sometimes hopeless efforts to reach a certain level of completeness or to define methods and tools to address the most demanding section of the lexicon, individual lexicographers, in terms of tools, expertise but also in terms of availability of thousands of legally/freely usable lexicographical descriptions, could work more cooperatively to both enhance the capacity of dictionaries to play a significant role in the preservation of a language and minimize the cost of each dictionary.

This of course supposes to some extent a different organization of lexicographical work, but it is clear that either at a national or international level, much could be done in a rather short time by sharing resources and content that could benefit hundreds of languages at once.

References

Cristinoi, A. & Nemo, F. 2013. Challenges in endangered language lexicography. *Lexicography and Dictionaries in the Information Age*, 126–132. Denpassar: Airlangga University Press.

Green, D. & Green, H. 2010. *Yuwit Kawihka Dicionário Palikúr – Português*. Belem : SIL.

Grenand, F. 1989. *Dictionnaire Wayãpi-Français, Lexique Français-Wayãpi*. Paris: Peeters/Selaf.

Grenand, F. (ed.). 2009. *Encyclopédies Palikur, Wayana, Wayãpi: Langue, milieu et histoire, fascicule encyclopédie des Amérindiens de Guyane*. Paris: PUO-CTHS.

Grice, P. 1975. Logic and conversation. In *Syntax and Semantics, 3: Speech Acts*. P. Cole & J. Morgan (eds), 41–58. New York NY: Academic Press.

Hartmann, R.R.K. 2003. *Lexicography: Critical Concepts*. New York NY: Routledge.

Klimpp, D.A. 1995. *Vocabulario Piapoco-Español*. Bogotá: Asociación Instituto Lingüístico de Verano.

Launey, M. 2003. *Awna Parikwaki. Introduction à la Langue Palikur de Guyane et de l'Amapa*. Paris: IRD Éditions.

Mosel, U. 2011. Lexicography in endangered language communities. In *The Handbook of Endangered Languages*, P.K. Austin & J. Sallabank (eds), 337–353. Cambridge: CUP. doi:10.1017/CBO9780511975981.017

Schauer, J.G., Schauer, S., Yukuna, E. & Yukuna, W. 2005. *Meke Kemakánaka Puráka²aloji: wapura²akó chu, eyá Karíwana chu (Diccionario Bilingüe: Yukuna – Español; Español – Yukuna)*. Bogota: Editorial Fundación para el Desarollo de los Pueblos Marginados.

Urquía Sebastián, R. & Urquía Sebastián, W. 2008. *Diccionario yine–castellano (parcial)*. <http://www.lengamer.org/publicaciones/diccionarios/Dic_Prelim_Yine.pdf>

Jewish language varieties

Loss and survival

Bernard Spolsky

University of Bar-Ilan

Created in the special circumstances of diaspora isolation, persecution, and regular migration, Jewish language varieties that survived the Holocaust have proved, with the exception of Hasidic Yiddish, to be fragile and are becoming extinct as spoken vernaculars. Emancipation, especially when accompanied by admission to public and state schools, generally led to language shift. But postvernacularity, the building of a metalinguistic community through the preservation of the symbolic value of a language in place of its communicative use, has enabled many Jewish groups to maintain the identity associated with their ethnic heritage.

Keywords: diaspora, ethnic heritage, identity, Jewish, loss, migration, shift, postvernacular

1. Fragility and loss of Jewish language varieties

In the introduction to a book of Yiddish stories by the Soviet writer David Bergelson, Werman (1996: xv) describes a memorial meeting she attended in Jerusalem in 1993 for Bergelson and twelve other Yiddish poets murdered on Stalin's orders on August 12, 1952.

> The program began with a girls' choir singing a medley of Yiddish songs, followed by the actress reading one of Hofstein's Yiddish poems. And that ended the sound of Yiddish for the evening – no one else uttered a Yiddish word. The master of ceremonies introduced the speakers in Russian, the speakers addressed the audience in Russian, all the announcements were in Russian, and all the questions from the audience were in Russian. Yiddish, it seems, was reserved for the ceremonial rites: little girls crooning *tumba-la-laika* and aging actresses declaiming Yiddish verses, nostalgic concessions to a remembered past.

An objective review of the current state of Jewish language varieties would find almost all of them in similar straits, extinct or threatened and kept alive by elderly surviving native speakers and the nostalgia of a small group of enthusiasts.[1] Except for a few cases – Yiddish among followers of some Hasidic dynasties (Katz 2004), Juhuri (Shapiro 2010) spoken by Mountain Jews in Israel, some dialects of Judeo-Aramaic (Y. Sabar 2003), and two questionable[2] varieties – Judeo-Georgian (Enoch 2013) and the Marathi of Bene Israel (Weil 2005) – what remains are the fading memories and usage of a few elderly speakers, some written works read by scholars, and a few devotees who are willing to learn but not speak the variety as an everyday language. Ignoring its continuity among followers of some Hasidic dynasties, many people consider even Yiddish to be seriously endangered, and some scholars refer to it as in a "postvernacular" state (Shandler 2006) or as forming the basis of "a metalinguistic community" (Avineri 2012). This status seems to apply also to other surviving varieties like Judeo-Greek (Krivoruchko 2011) and Judeo-Arabic. Some varieties are long gone such as the Judeo-Czech suspected by Jakobson and Halle (1964, 1985) and M. Weinreich (2008) to be the language of the early glosses on medieval Hebrew manuscripts but argued by Czech linguists to be Old Czech written in Hebrew letters (Uličná & Polakovič 2013), or barely survived into the 20th century like Western Yiddish (Jochnowitz 2010; Starck 1994, 2007) or Jewish Malayalam (Gamliel 2013) or Judeo-Spanish[3] (Bunis 2010).

Why did these varieties prove to be so transitory, lacking the loyalty needed to maintain family and vernacular use? First, they developed in very special conditions, in Jewish communities locked or isolating themselves in real or virtual[4] ghettos or mellahs. After the destruction of the Second Temple by the Romans

---

1.　An earlier version of this paper, focused on the loss of Jewish language varieties, was given at the 2014 LAUD conference, and a modified form published as Spolsky (2015). Thinking over discussions at the conference about the nature of language endangerment, I started to wonder about Salikoko Mufwene's notion that language shift should be seen as normal evolutionary adaptation (Mufwene 2001). This paper is then the next stage, which is still being rethought. I am grateful to a reader of the manuscript of this paper for some suggested changes and additions.

2.　Generally considered dialects rather than distinct varieties.

3.　There is disagreement about the names of Jewish varieties (Bunis 2008). Some scholars argue for Judezmo or Dzhudezmo for the spoken variety of Jewish Spanish that developed in the Balkans and Turkey, preferring to keep Ladino for the written form. Another popular name is Spaniolit.

4.　The *shtetlakh* (small towns) of Eastern Europe had a mixed Jewish-gentile population, but the two communities were segregated and they lived separate lives. The ghettos and *Judengasse* of Europe and the mellahs of North Africa were sections of the cities where Jews lived, in minimal contact with the non-Jewish majority.

and the expulsion from their homeland, Jews spread throughout the Diaspora where a sizeable number, perhaps a sixth,[5] maintained their religious and ethnic identity. Like most migrants, Jews tended to live close to each other, a practice strengthened by religious requirements forbidding them to ride on the Sabbath which made it desirable to live within walking distance of a synagogue; they also observed dietary restrictions that discouraged contact with others. As a result, Jews commonly lived in recognizable quarters. Strengthening this segregation, both Christians and Moslems had rules against close association with Jews, under the Pact of Omar[6] (Stillman 1979, 1991), Jewish women were not permitted to use the same bathhouses as Moslem, and both Christians and Moslems had rules against Jews hiring Christian or Moslem servants. In many cases, the Jewish quarters became walled areas that could be defended against repeated attacks by anti-Jewish mobs, urged on by Christian or Muslim clergy. Later, on the model of the Ghetto of Venice, the Jewish quarters were closed off by local governments, and Jews were forced to live in them.[7]

As a result of these measures, Jewish migrants had limited contact with the host gentile community, a situation that encouraged preservation of the languages they brought with them. Only at rare times of tolerance, such as Moslem Spain and pre-Crusader Europe, was there sufficient interaction for anyone other than those doing business with the outside world to achieve fluency in the local language. In addition, the existence of separate educational systems kept Jewish children apart, while at the same time enabling them to gain the literacy that was vital to professional and commercial skills (Botticini & Eckstein 2012) as well as preserving knowledge of Hebrew.

In these circumstances, Jewish language varieties were created by the fusion of an earlier vernacular (Jewish varieties of Aramaic, Greek, Latin, Arabic for instance, brought from countries of earlier residence) with the sacred and literate proficiency in Hebrew and Talmudic Aramaic taught to all Jewish males, and the varied contacts of Jewish professionals and traders and their wives with speakers of the local coterritorial non-Jewish language. A move to a new sociolinguistic environment, either as a result of periodic expulsions by Islamic or Christian rulers, or of voluntary migration seeking new trade and professional opportunities, led

---

5. The Jewish population is estimated to have dropped from about six million in the first century CE to one million in the eighth.

6. Instituted in the 7th century CE, these rules were modified and variously implemented and continue where Jews remain in Arab lands.

7. Before the Ghetto in Venice was established in 1516, Jewish traders had to leave the island at nightfall (Calimani 1988).

to a change in the sociolinguistic situation and the coterritorial language and was naturally accompanied by the evolutionary development of a new variety.[8] When there was isolation, including the more or less voluntary segregation of Jews and gentiles in the *shtetlakh* of Eastern Europe, and the religiously driven insularity of present-day Hasidim, there could be language maintenance.

2. Shift to the standard

Emancipation and secularization in the modern world produced a new set of conditions: after the French revolution, there were fewer locked communities in Europe,[9] there was freer contact with higher status non-Jewish languages, and the compulsory state secular education led to Jews learning and speaking a standard language, whether Hebrew, German, English, Russian, French, Spanish, Turkish, Persian or whatever, although a Jewish accent may have shown how recently they immigrated. Under these changed conditions, the Jewish language varieties began to be lost after two or three generations of emancipation or emigration to a more tolerant society.

The major exception was Yiddish. Western Yiddish, considered to be the first branch to develop, was preserved in written form in a goodly number of translations of secular romances. It started to be lost as a vernacular with the Enlightenment in 18th century Germany through absorption into the standard variety of German to which its speakers easily switched, although it was still alive in the 19th century: M. Weinreich (2008: 722) notes that the German Jews who migrated to the USA between 1830 and 1870 still spoke it. However, it was no longer expanding and soon died out in Germany (Hutterer 1969); there were vestigial uses in Alsace (Starck 1994, 2007) and Switzerland (Fleischer 2005) where it was kept up especially by horse traders who increased the number of Hebrew words they used in order to keep their conversation unintelligible to German customers (Guggenheim-Grünberg 1954). In both these areas, which were highly multilingual with distinct local varieties, preservation was easier until quite recently, but there are no longer speakers.

Eastern Yiddish however flourished as the vernacular of most Jews and later the written variety of a growing Ashkenazic secular culture (M. Weinreich 2008).

---

8. This notion of evolutionary shift of language varieties is expressed by Mufwene (2001, 2005). It agrees with the proposal that we should focus on speakers rather than languages (Labov 2008).

9. In Eastern Europe, until the Revolution Jews were restricted to the Pale of Settlement (Klier 2010).

As long as Jews were restricted to *shtetlakh* in the Pale of Settlement, the small towns which they shared with Slavic-speaking gentiles, providing useful services as innkeepers and traders, their internal literary culture was perhaps limited. Published religious literature was mainly in Hebrew, though an important body of writing started to appear in Yiddish, nominally for "women and illiterate men" (M. Weinreich 2008).

In the early 20th century, however, there were major demographic changes as Jews were permitted to move to larger cities. Yiddish culture reached its zenith in the period between the world wars in three major urban centers, New York, Warsaw and Moscow, each with a population of hundreds of thousands of Jews.[10] This provided the critical mass that was able to support a high Yiddish culture. There were three key positive elements: a large population of Jews for whom Yiddish was the vernacular, speaking the language in everyday life; a strong and growing number of newspapers and magazines (in which books could be published as serials); and a number of Yiddish theatres producing plays regularly. In 1936, there were some 27 daily newspapers and a hundred weeklies published in Poland in Yiddish (U. Weinreich 1949: 165). In 1930, the circulation of the Yiddish daily press in the United States was over half a million.

D. E. Fishman (2005) is fully justified then in agreeing that this was a period of growing strength in Yiddish literature and culture. The major Jewish centers in Moscow, Warsaw and New York were newly formed immigrant communities – in New York, large numbers of Jews started arriving only at the end of the 19th century, a process that was slowed down after the 1927 Immigration Act discriminated against migrants from anywhere except Northern Europe; Moscow had been closed to all but a handful of rich elite Jews until the Russian revolution in 1917; and Warsaw too developed as a Jewish city only in the 20th century. In each, Yiddish high culture depended on a small number of activists and writers, while many Jews were busy assimilating into the general community, sending their children to government public schools which used English in New York, Russian in Moscow, and Polish in Warsaw.

In Warsaw,[11] there was a large number of Jewish schools, supported for a while by government funding that was reduced as a result of developing anti-Semitic attitudes and policies. It appears that Jewish girls, better educated than the general population, were more likely to attend state schools and to shift to Polish, which they taught to their children (Bacon 2009). The increasing separation of

---

10. There were other centers; Vilna considered itself important, with YIVO (Kuznitz 2014) located there, but it only had about 60,000 Jews.

11. K. Weiser (2015) explains why Warsaw became the center of the Yiddish movement rather than Vilna or Moscow.

the communities is shown by the fact that in the 1921 Polish census, about 75% of Jews claimed Yiddish as their native language, while 80% claimed it in the 1931 census.[12] Increasing numbers of Jews were attending state schools. In 1934–1935, about 250,000 Jewish students were enrolled in Jewish schools and 480,000 in non-Jewish state, municipal and private schools. 80,000 were in Agudas Yisroel *kheyders* and yeshivas, another 48,000 were in private *kheyders* and 20,000 girls in Beth Jacob schools. Thus, nearly twice as many were in schools that used Polish rather than Yiddish or Hebrew.[13] There had been a strong movement for secular Yiddish schools, supported by the Folkparty, led by Noah Prylucki, who from 1918 until he moved to Vilna after the German conquest, led a movement which tried to challenge the Hebrew-based Zionist school movement (K. I. Weiser 2011). The party never gained popular success, and Prylucki and other leaders were killed by the Nazis. Most of the Polish community was wiped out by the Holocaust, so that speculation about a non-existent future is inappropriate.

In the Soviet Union, a strong government policy developed to remove Judaism from the Jewish schools, leading to loss of Hebrew religious terms in Soviet Yiddish and an increase in Russian borrowings (Estraikh 1999). The number of Jewish schools in the Soviet Union was reduced in 1938, by which time the percentage of Jews claiming Yiddish as their first language had fallen from 73% in 1926 to 40% (Estraikh 1999). Hebrew was banned, but Yiddish continued to be recognized and a Jewish Anti-Fascist Committee was established to seek international Jewish support for the war against Germany. In 1948, the committee was disbanded and in 1952 thirteen Yiddish poets and writers were arrested and executed after a closed trial. The Soviet Jewish community shifted to Russian.

In New York a network of religious schools continued, the Hasidic schools among them favoring Yiddish, and there was until 1960 a number of secular Yiddish schools (Margolis 2013; Parker 1978; Yefroikin 1955). Most Jews however were educated in public schools; non-Hasidic Jewish schools also largely used English as the language of instruction. The results can be seen in the 2011 US Census (American Community Survey), where there were 162, 511 adults claiming Yiddish as a first language (less than half of the 1980 figure and fewer than the 216,615 claiming Hebrew).

It was possible then that this was a one-generation burst of Yiddish culture, but its stability was not to be tested: in Eastern Europe, after 1939, the Nazis wiped out many of the remaining Yiddish speakers. In the Soviet Union, recognition

---

12. Over 10% claimed Hebrew, reflecting commitment to the Hebrew school movement that was developing in interwar Poland.

13. There figures come from Table 12 in Fishman (1991: 402–203).

of Yiddish in the 1920s was reversed by the banning of Yiddish and the arrest and murder of the poets (Shneer 2004). As early as 1922, there had been strong attacks on the Jewish religious schools, in the Communist Yiddish publications which aimed to secularize and Russianize Jewish culture (Shternshis 2006). Even without this, strong pressure for Russian in the school system worked to endanger and weaken Yiddish. So the Soviet encouragement of Yiddish as a minority language continued only as a symbol in Birobidzan. In Moscow and the rest of Russia, Yiddish continued only as the language of elderly Jews and a tiny group who maintained religious observance.

Attempts to establish Yiddish in Israel failed as a result of the growing ideological institutional power of Hebrew. In 1928, a small group of Yiddish writers, members of the left wing Poelei Zion movement (whose 1906 decision to use Hebrew rather than Yiddish in its publication had been a turning point in the revival of Hebrew), started a magazine that was intended to form a base for the masses of Yiddish speaking immigrants that they expected who would swamp Hebrew speakers. They did not arrive, partly because European Jews were not ready to go to Israel (the majority of religious leaders was anti-Zionist) and partly because of the British restrictions on Jewish immigration to Palestine. When the survivors finally did come after the establishment of the State, the Yiddish speakers from East Europe were balanced by a large immigration of Jews expelled from Arab lands, non-speakers of Yiddish, so that Hebrew easily maintained its position as not just official but also a needed lingua franca (Spolsky & Shohamy 1999).

Yiddish speaking Jews in the Soviet Union had been the first target of the invading Nazi army, which wiped out the Yiddish speakers of Poland and other Eastern European countries. Jews who survived or returned to Poland after the war suffered a new set of anti-Jewish activities in 1944–1946, leading to further emigration; thus, of 300,000 Jews in Warsaw before the war, there are now between 5000 and 20,000 Jews in the country as a whole, and Yiddish is only spoken by an elderly remnant, although there are reports of some cultural revival; a Jewish State Theatre continues in Warsaw.

Those Yiddish speakers who moved west and continued to the Americas also came under external pressure. In Belgium, a small ultra-orthodox group maintained separation and Yiddish around the diamond industry but by now most are also proficient in French, Dutch, and Hebrew. In England, Yiddish speaking refugees were urged by their leaders to learn the English needed for economic success and social acceptance: only in Stamford Hill does a small block of Hasidim remain some of whose sons are introduced to Yiddish when they begin religious education. The pattern in the USA was similar; in New York, the Yiddish theatres and the Yiddish newspapers did not survive the passing of the immigrant generation.

Nazi genocide and Soviet anti-Jewish activity wiped out the large Eastern Yiddish speech communities, leaving a small group of activists studying the modern literature and mourning the loss of the language, which has been adopted as a new means of closing off an ultra-orthodox Jewish life in some Hasidic sects (especially Satmar) whose males acquire it when they start to attend religious schools at the age of six. In all, Yiddish between the two world wars does illustrate a key feature of Jewish languages: their fragility even when strong.[14]

In the case of Jewish languages other than Yiddish, all lacking the institutional support of governments (token recognition of Yiddish by the Israeli government and some Scandinavian countries is virtually meaningless), a strong school system (provided for some Hasidic boys but not for any other Jewish language apart from Hebrew), and a vibrant literature, the languages may soon be of interest only to scholars and enthusiasts. There is perhaps a lesson for other endangered language activists: full maintenance of a language depends on a consensual ideology (belief in the value of the variety), encouragement of practice (provision of opportunities for use whether literary or in other media, and in important social functions including especially education), and continued efforts to achieve recognition by the community and government. In a set of essays on the sociology of Yiddish, Fishman (1991: 310) notes that the only upward trend in indicators is the number of tertiary level courses; however, he argues throughout that it still lives and may well be further restored.

In the absence of institutional support and educational programs, Jewish languages depended for survival on the continued immigration of speakers (as is the case with Spanish in the US) or continued segregation imposed either internally (as with the Amish and the Hasidim) or externally (as in the ghettos and with US Blacks). But Jewish education, even before it was replaced after Emancipation by public schooling in the dominant state language, was strongly weighted towards the Hebrew-Aramaic of liturgy and sacred texts rather than the Jewish varieties in which it was often taught.

Lacking the status of Hebrew-Aramaic and the modern intellectual elite supporting Yiddish in Eastern Europe and the Americas in the first half of the 20th century, the other Jewish varieties have been easy prey to the standard varieties introduced through schooling. One such force was towards French, in North Africa and the Middle East, the result of language diffusion efforts conducted by the Alliance Universelle Israélite, a Jewish supported but government encouraged organization which established French medium schooling for Jewish children in

---

14. I am grateful to a talk by Avraham Novoshterm on 16 February 2014 at the opening of the Yiddish Winter Course at (ironically) Beit Ben Yehuda in Jerusalem given (another irony) in English.

many schools in the Mediterranean region. This was reinforced by the government schools in North Africa under French rule; Muslims were reluctant to attend, but Jews and Christians took advantage of them and sent their children to schools that promised educational and economic advantage. The shift to French led to the weakening of both Hakétia, the Moroccan variety of Judeo-Spanish, and the various forms of Judeo-Arabic that had served not just as internal community languages, but that had produced an important religious literature, both of sacred poetry that formed an important part of synagogue ritual and for the writing of philosophical work and religious commentaries that marked the Jewish contribution to the Middle Arabic used by non-Muslims (Hary 1995, 2009). Both Judeo-Arabic and Hakétia came up against the pressure of Hebrew hegemony when the Jews in Moslem countries were expelled and moved to Israel; similarly, those North African Jews who moved to France continued their growing commitment to the higher status French language.

The independence of Greece and Bulgaria changed the status of the Judeo-Spanish speaking Jews of the Ottoman empire who had been protected by the tolerance of the *millet* system that recognized minorities. Later, Greek and Bulgarian nationalism led to the loss of the widespread Judeo-Spanish which had developed there after the expulsion from Spain. When Salonika came under Greek rule, the large Jewish population needed to add Greek to their repertoire and the role of Judeo-Spanish was diminished (Naar 2011); there was some migration to Israel and the USA, and most of the remaining Jewish population was killed at Auschwitz by the Nazis. In Turkey, like North Africa, the French introduced by the Alliance schools became a language of the Jews, weakening the Judeo-Spanish that had been brought from Spain and that was starting to build a modern variety. By the 1920s, French was well established in the Turkish Jewish community and threatening Judeo-Spanish (Sarhon 2011). Both French and Judeo-Spanish were to be replaced by the Turkish promulgated and enforced by Kemal Ataturk after the revolution (G. Lewis 1999). The process of shift is confirmed by the recent publication of a Jewish bible and prayer book with translation into Turkish. But there remains a nostalgia for Judeo-Spanish that shows up in various postvernacular activities (Kushner-Bishop 2004a, 2004b) in Israel and the USA.[15]

Judeo-Arabic varieties, developed by Jews living in the countries that had been conquered by Islam, also showed vulnerable language loyalty. Jews in those Arab countries where they were permitted to live (Saudi Arabia was early declared to be only for Moslems) were tolerated along with Christians and Zoroastrians, but only in the status of *dhimmitude* (Stillman 1979, 1991; Ye'or 1985), governed

---

15. But the 2011 Community Survey in the US reports only 136 speakers of Ladino; the figure is so low presumably they reported their language as Spanish.

by unevenly imposed rules that kept them as second class subjects, set head taxes, forbade them to employ Moslem servants and their women to use the same bathhouses as Moslems, and restricted their access to the classical language of the Qur'an.[16] Jews in these countries maintained their own educational systems, which taught boys Hebrew sacred texts through vernacular Judeo-Arabic. When the choice of education in French was offered in 19th century North Africa, it was quickly taken up. When Jews were expelled from Arab countries after the United Nations decision in 1947 to partition Palestine, those who went to France shifted to French and those who immigrated to Israel soon shifted to Hebrew, albeit for a while a marked Sephardic variety.

Both Turkey and Persia, while being conquered and ultimately converted by Islam, had not become Arabic speaking. The Jews, who had lived in Persia since the Babylonian Exile, spoke a number of different varieties (Shapira 2003). A strong Judeo-Persian literature developed over the centuries (Fischel 1960; Spicehandler 1968) and local spoken Jewish varieties were still preserved by older immigrants to Israel from the smaller centers, but younger and better educated Persian Jews had developed control of Farsi, the national language, even before emigration.

There were also Jewish varieties of other Iranian languages, including Yazdi and Kurdish, but these too were lost once there was access to public education and mass media. Kaganovitch (2008) reports on the Bukharan Diaspora; and the Endangered Language Alliance is studying Bukhori maintenance in the large community in Queens New York, finding that older members still speak the language regularly, but few young people use it. In the 1920s, the Soviets had encouraged activity in Bukhori and Judeo-Tajik, but this policy was not continued and Russian was promoted. Also in the 1920s, there was Soviet support for publications and other activities in Judeo-Tajik, but school use and publication were banned in 1940, with a consequent shift to Tajik, Russian and Uzbek; after immigration to Israel in the 1970s, some cultural activities in Judeo-Tajik re-emerged. In Israel and Russia, there is a survival of Judeo-Tat, the language of the Mountain Jews, also known as Juhuri, but younger speakers are reported to be shifting to Russian and to Hebrew.[17]

---

16. There were exceptions, as in Spain before the Christian reconquest, and more oppressive measures, as the forbidding of riding camels and horses in Yemen which lasted as long as Jews were allowed to live there.

17. "At present there are about 100,000 Caucasian Jews living in Israel, most of them in Acre, Haifa, Be'er Sheva, Hadera, or Akiva, Netanya, Ashdod and Ashkelon", reports Litman (2010). She cites Vitaly Shalem, who wrote an MA thesis on Mountain Jews, saying that only in northern Azerbaijan are there children speaking Juhuri to each other; she also quotes Boris

In other Islamic countries where Arabic did not replace other languages, varieties of Aramaic continued to be spoken by Christians (800,000[18]) and Jews (perhaps 25,000), though in Iraq where roads allowed easy access, there was a shift to Arabic (Y. Sabar 2003). Zakho was an important city, and Jews from there kept up the variety alongside Kurdish for communication with non-Jews, Hebrew for religious purposes, and Russian for schooling (A. Sabar 2008). But most of the Jewish dialects are now extinct or virtually so: Barzani Jewish Neo-Aramaic, Lishana Deni, Lishana Noshan, Lishan Didan. There are said to still be speakers of Hulaulà in Israel (Hezy Mutzafi, personal communication).

In Italy, not united linguistically until the 20th century with dialects still important (Lepschy & Lepschy 1998), Jewish language varieties were derived from the regional dialects (Jochnowitz 2013). There was a variety of Judeo-Piedmontese whose final stages are recorded by Primo Levi (Jochnowitz 1981; George Jochnowitz Levi 1985). Jochnowitz (n.d.) refers to a Judeo-Mantuan dialect. There was a Jewish variety of Venetian (Fortis & Zolli 1979), and a Florentine variety (Jochnowitz 1978). Probably the only surviving variety is the Roman one, Giudaico-Romanesco, still being used in the 1980s through the encouragement of a theatre group of young adults (de Benedetti 1997). Here too, we have evidence of a nostalgic vestigial use of a Jewish variety.

The variety spoken by Georgian Jews is a border-line case, the result of a 20th century debate in the Jewish community over identity. Three points of view emerged: one position was taken by traditional rabbinical leadership who were afraid that emancipation could destroy the centrality of the synagogue; a second was that of the assimilators who called for Jews to be Georgians with a Jewish faith; and the third, led by the Zionists, who wanted to integrate Georgian Jews with international Jewry, stressed Jewish identity and the place of Israel. That the Zionists won out is shown by the immigration of the largest proportion of Georgian Jews to Israel in the 1970s. Since the 1980s, all this has been expressed in a widespread belief that Georgian-Jewish relations were always friendly and warm (Kakitelashvili 2012). Linguistically, this view is supported by the closeness of the Jewish variety to the standard. Wikipedia claims that there are about 85,000 speakers of "Judaeo-Georgian" including 20,000 speakers in Georgia (a 1995 estimate), and about 59,800 speakers in Israel (a 2000 estimate); it also reports approximately 4,000 speakers in New York and undetermined numbers in Russia, Belgium, Canada and elsewhere in the USA. *Ethnologue* (M. P. Lewis, Simons & Fennig

Hanukayev, who still writes in the language, reporting that few Israeli Mountain Jews speak Juhuri to their children.

18. Before the attacks by the Islamic State currently taking place in northern Iraq (Perlin 2014), killing or driving out Aramaic-speaking Christians.

2013) is more cautious, echoing Friedman (2010) in suggesting that it is perhaps just Georgian with some added Hebrew lexicon. Friedman (2010: 118) considers it "an ethnolect of Georgian like the English of American Jews for English." Moskovich and Ben Oren (1981) conclude that "while it is widely held that there is no Judeo-Georgian language, and Georgian Jews themselves claim they speak Georgian, there are many Hebrew and Aramaic loanwords and calques in three levels of Judeo-Georgian speech: everyday communication (less among younger speakers), Bible and Talmud interpretation, and argot." Enoch (2013) has studied the Hebrew components of Judeo-Georgian.

Asian Jewish varieties have also largely disappeared. In India, Jews in Cochin formed two communities, maintaining the separation between castes practiced in Kerala and elsewhere in India. The Malabari (Black) Jews claim to be descended from traders who arrived at the time of Solomon.[19] The Paradesi (White) Jews arrived in the 16th century, after the expulsion from Spain. Both groups spoke a variety of Jewish Malayalam different from (but mutually intelligible with) the language of Kerala, but the variety seems to be no longer spoken by those Jews who remain in Kerala and to be virtually extinct among those who emigrated to Israel between 1950 and 1970 (Gamliel 2013).

3. Survival through postvernacular practices in a metalinguistic community

This rounds out the story of loss, but I have already hinted at cases of partial survival. If we treat Georgian as not being a Jewish variety, then it fits with all the other immigrant languages (Russian, French, English, Amharic, and so on) brought to Israel. Their loss is part of a study of immigrant language shift and the existence of the language in its homeland is not threatened by what happens in any one diaspora. It is rather the recognized Jewish language varieties like Yiddish and Ladino and Juhuri and Judeo-Aramaic that are our topic.

The earlier part of this paper has shown the fragility of Jewish varieties, the low level of loyalty, and the ease with which they disappear in changing environments. There are two counterforces that we need to note. The first is the case of some ultra-orthodox sects of Hasidim who have established an educational system mainly for boys that guarantees their proficiency in Yiddish and enables them to maintain isolation not just from the non-Jewish world, building their own voluntary ghetto and sharing the special dress and food laws that keep them separate, but also from other Jews with more modern ways of life. The second is

---

19. For discussion of Black Jews, see Spolsky (2014: 92–93).

the recent growth of what Shandler labels *postvernacularity* as some Jews choose various activities (courses, clubs, web lists, theater, camps) to maintain and build endangered heritage languages and varieties. What the two may have in common is that they can manage without natural intergenerational transmission because they provide other methods of teaching or learning languages; what divides them is that the Hasidim aim at daily vernacular language use, keeping a special status for Hebrew-Aramaic as language of study and prayer, a status which the secular Yiddishists assign to their heritage language,[20] using the common local vernacular for daily life.

The concept "postvernacular" is usually attributed to Shandler's study of "Yiddishland", but there seems to have been earlier uses: Preston (2004: 153) uses it to mean the learning of any language after the first. Coining the term for the situation of Yiddish is also claimed by Kuznitz (2004).[21] Shandler (2006: 4) defines postvernacularity as "having an affective or ideological relationship with Yiddish without having command of the language." The same phenomenon is described by Avineri (2012, 2014) as the creation of a "metalinguistic community which talks about a language rather than speaking it." Applying this distinction clearly excludes Hasidic vernacular proficiency and use, but if we employ Preston's definition of a language acquired other than as a first language, we might consider the teaching of Yiddish to boys when they start *heder* and to girls in Yiddish classes in school to include a proportion of Hasidic Yiddish in the same category.[22] Proficiency and daily use however suggest it is better to keep secular and Hasidic Yiddish apart.

That being the case, what then are the features of postvernacular status that implement the "affective or ideological relationship"? Fundamental, Shandler (2006: 4) says, is that Yiddish moves from earlier instrumental and communicative use to a secondary, meta-level status with symbolic value replacing semantic. Postvernacularity, signs of which emerged even before World War II, means that Yiddish hasn't been lost, but its status in the Jewish linguistic repertoire has been changed. It has had to deal with the challenge of comparison originally with German and now with Hebrew and English, each standing not just for the outside world but also representing modernity and a higher culture. Yiddish postvernacularity is now associated with an imagined Yiddishland, a recreated idealized virtual *shtetl*, where the language substitutes for the homeland.

---

20. Fishman (2002) argues that since the Shoah, Yiddish is a holy language.

21. Margolis (2011) attributes it to both Shandler and Kuznitz.

22. Some Hasidim speak Yiddish to their babies, more outside than inside Israel (Fader 2001). Katz (2004) believes that the numbers are high enough to guarantee the future of Hasidic Yiddish.

Shandler lists some of the activities that implement this ideology. First is the teaching of the language, which for generations was unnecessary because Jews spoke it from childhood. Yiddish language schools developed, he says, only during and after the First World War, in Poland and Lithuania mainly, where they competed with Zionist schools which taught in Hebrew and the state schools which used the national languages. With state recognition, Yiddish schools received government support in the Soviet Union in the 1920s and 1930s (Shneer 2004). But this ended with the Second World War and the Holocaust.

Starting in the late 19th century, two million Jews had emigrated from Eastern Europe to the United States, where most were persuaded to shift to English. Some religious education continued in Yiddish, expanded after the arrival of Satmar and other Hasidic communities after the war. A network of secular Yiddish schools also existed until 1960 (Freidenreich 2010), and Yiddish was introduced into university studies, especially with the publication of *College English* by U. Weinreich (1949). This teaching has continued with the offering of many Yiddish language courses for adults, but there has been nothing like the early childhood program that played a major part in the regeneration of Māori in New Zealand or the intensive Hebrew *ulpan* for new immigrants in Israel. Avineri (2012) calls this continuing teaching "nostalgia socialization". Postvernacularity now means teaching mainly adults, and aiming for some knowledge but not for daily use.

Another feature is the borrowing of Hebrew and Yiddish words and phrases into the new host language. This may be a sign that speakers are new immigrants learning a new language, code-switching and borrowing terms they do not know in the language, a normal process in language shift. But Benor (2012) has drawn attention to a reverse of this process, as fluent native speakers of English (and probably French and Spanish too) add Hebrew-Aramaic and Yiddish terms and expressions to their speech in order to show their status as newly-observant Jews, signaling their growing identity with speakers of Yiddish without actually speaking it. This too is surely a feature of postvernacularity, a kind of limited learning with important symbolic meaning.[23] In contrast, the continuing vernacular use of Hasidic Yiddish is marked by borrowing from English. Katz (2004: 380) sees this as a sign of life, and notes that it is one of the complaints of modern secular Yiddish scholars about the variety.

The second feature that Shandler mentions is translation, a well established Jewish practice: the Talmud interprets a phrase in the Book of Ezra as suggesting that the Targum in Aramaic began to be added to the synagogue ritual soon after the return from the Babylonian Exile. For Yiddish, it may be exemplified by the

---

23. Benor suggests that it may be the way in which a new variety of Jewish English is being created.

Tsene-rene, first published about 1600 and still in print, a collection of Bible translations and commentaries in Yiddish intended for women. In the 19th and early 20th century, while there was still a large Yiddish speaking and reading public, there was also a good deal of translation of world literature into Yiddish.[24] In the 20th century, however, translation started in the opposite direction, with publication of many works of Yiddish literature in English or Hebrew or other languages more accessible to the postvernacular readership.[25]

The next feature of postvernacularity that Shandler describes he calls "performance art", which ranges from two academics speaking Yiddish to each other loudly at a conference to the *yidish-vokh* (Yiddish week) that Mordke Schaechter initiated in 1975, a week long Yiddish family immersion retreat conducted every summer in the USA. Similar events are also sponsored by *Yugntruf*, a young adult association formed in 1964 in the US to encourage the maintenance of Yiddish. The fact that these activities are concentrated on adults rather than the kind of pre-school and primary immersion teaching that was the core of Māori revival activities (see for example Spolsky, 1987, 2005, 2010) and that is the core element in Hasidic Yiddish education is further evidence of postvernacularity. This is a good place to recall that YIVO was conceived as a university level research institute to cap the existing school programs, and that it continues in this higher role as a major supporter of postvernacular activity (Kuznitz 2014).

Theater is another area of Yiddish performativity. There were in fact a dozen or more Yiddish theaters in the USA up to the Second World War, and National Yiddish Theater *Folksbiene* founded in 1915 still exists in New York. In Israel, there is also a Yiddish theater company, *Yiddishspiel,* founded in1987, that gives regular performances of plays and musical comedy in Tel Aviv and other cities.[26] Shandler also makes reference (p. 55) to the growing number of websites that encourage use of Yiddish, one of the more prominent of which is the remaining US Yiddish newspaper, *Forvertz*, which illustrates the working of postvernacularity by not only having an English edition, but a Yiddish edition which provides both a translation and a dictionary look-up app for every Yiddish word.

The *Forvertz* on-line translations and the surtitles used by *Yiddishspiel* and sub-titles in movies demonstrate a central feature of postvernacularity – access to the language and its symbolic heritage identity is made possible even for those

---

24. There is no evidence for a book with the title "Shakespeare translated into Yiddish and improved".

25. Isaac Bashevis Singer won the Nobel Prize for Literature in 1978.

26. The website notes that thanks to a subvention from the Poalim Fund, "Yiddishspiel productions are surtitled with simultaneous translations in Hebrew and in Russia."

who are not fully proficient in it. It is a cheap but effective means of maintaining the value of a heritage language. At the same time, it can serve as an entrance to greater knowledge and use of the language: the immersion camps and the classes start at the beginner's level, but those who are committed and continue can reach higher ranges of mastery.

Use of the internet to support postvernacular activities has been recognized by a special issue of the journal *Language & Communication*.[27] Benor (2011) introduces the issue, asserting that the internet is a valuable way to preserve endangered languages. Sadan (2011) describes websites for Yiddish, suggesting that still more can be done. Ladino websites are discussed by Brink-Danan (2011), analyzing in particular how activists explore and debate the notion of what she calls "Ladinoland." In a detailed review of the postvernacular developments of the several varieties of Judeo-Greek, no longer spoken but still attracting feelings of nostalgia, Krivoruchko (2011) considers the situation among Jews in Greece, and the imagined Greek identities of those who have emigrated to Israel and the USA. E. M. Dean-Olmsted (2011, 2012) studying the Syrian Jewish community in Mexico City shows how uses and definitions of terms for sub-groups reflect changes in attitudes within the community.

There are reports of postvernacular activities for other Jewish language varieties. The two institutions established by the Israeli government, the National Authority for Yiddish Culture and the National Authority for Ladino essentially fit this pattern. There is a World Congress for Georgian Jews, but its website is in standard Georgian. There are two theater companies in Israel which present plays in some smaller cities in Judeo-Arabic (Henschke 2014). A Kol Israel radio station, *Reka*, broadcasts a daily program in Judeo-Moroccan. There are said to be monthly meetings in Jerusalem of Jews from Zakho speaking a variety of Judeo-Aramaic.

All these postvernacular practices serve to maintain language-related identity for the Jewish groups, a kind of *landsmanshaft* (hometown society) that provides spiritual support rather than the practical financial and social support of the New York based immigrant societies. They are a way for the third generation to reproduce (or reinvent) aspects of their grandparents' culture, using the language as a symbolic connection. True, it doesn't preserve the language as a functioning vernacular for linguists to analyze, but it does preserve identity for the metalinguistic community who see it as a cherished heritage.

---

27. Volume 31, number 2, edited by Sarah Benor and Tzvi Sadan.

4.   Hasidic Yiddish maintenance

It is only Yiddish that still has speakers among Hasidic sects committed to continued use of the language as a vernacular, although even among Hasidim there appears to be a division between males who use it from the time they begin *heder* at the age of 6, and the females who are more likely to be proficient in Hebrew in Israel and English in the US. Given the tendency of modern Hasidic dynasties to splinter, usually associated with a dispute over succession in leadership, and the closed nature of the Haredi world with its resistance to modernity including censuses and scientific study, it is hard to find precise details of numbers or language patterns. Based generally either in Israel or in the USA, there are also important Hasidic communities in London and Antwerp. Each group is known by the city of origin of its dynasty. The largest nine are Belz (mainly now in Israel, about 7,000 families), Bobov (headquartered in Borough Park, Brooklyn, about 10,000 followers), Lubavich (Chabad, its center in Crown Heights, Brooklyn, with a major settlement in Israel and missionaries spread through the world, with perhaps as many as 200,000 followers), Ger (located in Jerusalem, about 13,000 families), Karlin-Stolin (Israel), Sanz Klausenberg (divided between Borough Park and Kiriat Sanz in Israel), Satmar (in Kirias Joel and Williamsburg, New York, about 150,000 worldwide), Skvar (New Square, New York, about 20,000 followers), and Vizhnitz (New York and Bnei Brak, Israel, about 5000 followers). There are at least 28 other smaller dynasties, and many other tiny ones which survived the Holocaust. While they share many common features, they vary in details of practice, such as dress, liturgy, food, and language policy, making generalization difficult. But most share in recognizing the basic importance of Yiddish as a language for teaching sacred Hebrew-Aramaic texts and also for home and daily life, with education of boys in Yiddish especially important.

Given the complexity and the number of groups, there have been only a few studies with sufficient detail to summarize language patterns, but there is general consensus with the opinion of Katz (2004) that the Hasidic world is the main bastion of Yiddish vernacular continuity. As a general rule, most Hasidic communities contain a majority of adult speakers of Yiddish, the language commonly of their communal life. Some even use it writing, though the emphasis is on writing in Hebrew.

But it is more difficult to decide on the method of intergenerational transmission. Studies such as Baumel (2006) and Fader (2009) seem to suggest two complementary patterns. In one, noted more in the diaspora communities than in Israel, and more in the isolating sects like Satmar and rare in Lubavitcher, Yiddish is spoken in the family and to young children, demonstrating the "natural" transmission that language activists seek. In the USA, Satmar boys, it is reported,

do not meet English until it is taught them as a foreign language to be used to present the stigmatized but state-required secular studies. A common pattern is that Hasidic boys are brought up by mothers speaking the local language (English in the US and the UK, Hebrew in Israel) and meet Yiddish as the main language of instruction when they start *heder* at the age of six; most are fluent by the time they reach Yeshiva, but in the United Kingdom, Ger boys are reported to have difficulty with Yiddish even when they reach bar mitzvah age. In Hasidic boys' schools, religious subjects which are considered the most important component are taught in Yiddish (or *Leshon Hakoydesh* – a modified Hebrew-Aramaic), while any secular subjects may be taught in the local language. In most Hasidic girls' schools, on the other hand, teaching is in the local language, with some classes for Yiddish added to the curriculum.

Thus, Yiddish in the Hasidic world shares aspects of vernacularity (its widespread use as a daily language of home, school, and community, and some transmission at home to young children) with some aspects of postvernacularity, especially the dependence on school teaching of the language.

5.   Postvernacular practices as a goal for threatened varieties

This evidence of the existence of postvernacularity as an evolutionary development in the changed sociolinguistic environment that led to the loss of Jewish language varieties in the 20th century perhaps offers a clue to possible goals for other endangered languages. The example of Hasidic Yiddish shows that maintenance of sufficient proficiency for regular everyday use is possible, but it demands strict discipline and willingness to isolate the community from many aspects of modern life. But even in the case of Hasidic Yiddish, it appears that school teaching (a key feature of postvernacularity) is a major element in preservation; there is some natural intergenerational transmission, but it is perhaps not a necessary feature. Of course, in cases where an endangered language continues to be spoken to young children (e.g. Catalan, many South American Indian languages, Inuit, Northern Sami, Welsh in the north), a logical goal of reversing language shift is to protect maintenance by raising the status of the language to make it able to resist the pressure of the standard language. But where the language is no longer spoken in the family and to young children, the continuity of ethnic and group identity may well be served by accepting postvernacularity as a goal and not a regrettable defeat. After all, Hebrew was maintained in a postvernacular status for two millennia by an educational system. From this point of view, one wonders if the concern over the failure to persuade Māori to use their language in the home and to develop Māori language education much beyond the elementary school

is as serious a tragedy as some now claim (Bauer 2008; Harlow & Barbour 2014; May & Hill in press).

If we accept the suggestion of Mufwene (2001) that language shift is a normal evolutionary response to changes in sociolinguistic environment, as stronger languages continue to dominate weaker (de Swaan, 1998a, 1998b, 2001), then perhaps postvernacularity, the building of a metalinguistic community, offers a way to preserve the symbolic identity value of a beloved language (J. A. Fishman 1997).

References

Avineri, N. 2012. Heritage Language Socialization Practices in Secular Yiddish Educational Contexts: The Creation of a Metalinguistic Community. PhD dissertation, UCLA.

Avineri, N. 2014. Yiddish endangerment as phenomenological reality and discursive strategy: Crossing into the past and crossing out the present. *Language and Communication* 38: 18–32. doi:10.1016/j.langcom.2014.05.002

Bacon, G. 2009. *Jewish Women: A Comprehensive Historical Encyclopedia. Jewish Women's Archive*. Poland: Interwar.

Bauer, W. 2008. Is the health of Te Reo Maori improving? *Te Reo* 51: 33–73.

Baumel, S.D. 2006. *Sacred Speakers: Language and Culture Among the Haredim in Israel*. Oxford: Berghahn Books.

Benor, S.B. 2011. Jewish languages in the age of the internet: An introduction. *Language and Communication* 31(2): 95–98. doi:10.1016/j.langcom.2010.08.001

Benor, S.B. 2012. *Becoming Frum*. New Brunswick NJ: Rutgers University Press.

Botticini, M. & Eckstein, Z. 2012. *The Chosen Few: How Education Shaped Jewish History*, 70–1492. Princeton NJ: Princeton University Press.

Brink-Danan, M. 2011. The meaning of Ladino: The semiotics of an online speech community. *Language and Communication* 31(2): 107–118. doi:10.1016/j.langcom.2010.08.003

Bunis, D. 2008. The names of Jewish languages: A taxonomy. In *Il mio cuore è a oriente. Studi di linguistica storica, filologia e cultura ebraica dedicati a Maria Luisa Mayer*, F. Aspesi, V. Brugnatelli, A.L. Callow & C. Rosenzweig (eds), 415–433. Modena: Quaderni di "Acme".

Bunis, D. 2010. Judeo-Spanish. In *Encyclopedia of Jews in the Islamic World*, N.A. Stillman (ed.), 69–73. Leiden: Brill.

de Benedetti, J. 1997. Dabbera in Scionaccodesce (Speak Giudaico-Romancesco): Keeping the Jewish-Roman Dialect Alive. PhD dissertation, State University of New York. <http://search.proquest.com/dissertations/docview/304391916/previewPDF/1413260 A251158ACBB1/1?accountid=14483>

Calimani, R. 1988. *The Ghetto of Venice*. Milan: Rusconi.

de Swaan, A. 1998a. A political sociology of the world language system (1): The dynamics of language spread. *Language Problems and Language Planning* 22(1): 63–78. doi:10.1075/lplp.22.1.04swa

de Swaan, A. 1998b. A political sociology of the world language system (2): The unequal exchange of texts. *Language Problems and Language Planning* 22(2): 109–128. doi:10.1075/lplp.22.2.01swa

de Swaan, A. 2001. *Words of the World: The Global Language System*. Malden MA: Polity Press and Blackwell.

Dean-Olmsted, E.M. 2011. *Shamis, halebis* and *shajatos*: Labels and the dynamics of Syrian Jewishness in Mexico City. *Language & Communication* 31(2): 130–140. doi:10.1016/j.langcom.2010.08.005

Dean-Olmsted, E.M. 2012. Speaking Shami: Syrian Jewish Mexican Language Practices as Strategies of Integration and Legitimation. PhD dissertation, Indiana University.

Enoch, R. 2013. Judeo-Georgian, Hebrew component. In *Encyclopedia of Hebrew Language and Linguistics*, G. Khan (ed.). Leiden: Brill.

Estraikh, G. 1999. *Soviet Yiddish: Language Planning and Linguistic Development*. Oxford: OUP. doi:10.1093/acprof:oso/9780198184799.001.0001

Fader, A. 2001. Literacy, bilingualism, and gender in a Hasidic community. *Linguistics and Education* 12(3): 261–283. doi:10.1016/S0898-5898(01)00056-0

Fader, A. 2009. *Mitzvah Girls: Bringing Up the Next Generation of Hasidic Jews in Brooklyn*. Princeton NJ: Princeton University Press.

Fischel, W.J. 1960. Israel in Iran (A survey of Judeo-Persian literature). *The Jews – Their History, Culture and Religion* (New York, 1949), 827.

Fishman, D.E. 2005. *The Rise of Modern Yiddish Culture*. Pittsburgh PA: University of Pittsburgh Press.

Fishman, J.A. 1991. *Yiddish: Turning to Life*. Amsterdam: John Benjamins. doi:10.1075/z.49

Fishman, J.A. 1997. *In Praise of the Beloved Language: A Comparative Review of Positive Ethnolinguistic Consciousness*. Berlin: Mouton de Gruyter. doi:10.1515/9783110813241

Fishman, J.A. 2002. The holiness of Yiddish: Who says Yiddish is holy and why? *Language Policy* 1(2): 123–141. doi:10.1023/A:1016181810717

Fleischer, J. 2005. *Surbtaler und Hegauer Jiddisch. Tonaufnahmen und Texte zum Westjiddischen in der Schweiz und Südwestdeutschland*. Tübingen: Niemeyer. doi:10.1515/9783110935523

Fortis, U. & Zolli, P. 1979. *La Parlata Giudeo-Veneziana*. Roma: B. Carucci.

Freidenreich, F.-P. 2010. *Passionate Pioneers: The Story of Yiddish Secular Education in North America, 1910–1960*. Teaneck NJ: Holmes and Meier Publishers.

Friedman, V.A. 2010. Sociolinguistics in the Caucasus. In *Handbook of Sociolinguistics Around the World*, M.J. Ball (ed.), 127–138. Abingdon: Routledge.

Gamliel, O. 2013. Voices yet to be heard: On listening to the last speakers of Jewish Malayalam. *Journal of Jewish Languages* 1: 135–167. doi:10.1163/22134638-12340004

Guggenheim-Grünberg, F. 1954. Horse dealers' language of the Swiss Jews in Endingen and Lengnau. In *The Field of Yiddish*, U. Weinreich (ed.). New York: Publications of the Linguistic Circle of New York (now the International Linguistic Association).

Harlow, R. & Barbour, J. 2014. Māori in the 21st century: Climate change for a minority language? In *Language Ecology for the 21st Century: Linguistic Conflicts and Social Environments*, W. Vandenbusch, E.J. Håkon & P. Trudgill (eds), 241–266. Oslo: Novus Press.

Hary, B. 1995. Judeo-Arabic in its sociolinguistic setting. *Israel Oriental Studies* 15: 73–99.

Hary, B. 2009. *Translating Religion: Linguistic Analysis of Judeo-Arabic Sacred Texts from Egypt*. Leiden: Brill. doi:10.1163/ej.9789004173828.i-360

Henschke, J. 2014. Saving Jewish languages, *Haaretz*.

Hutterer, C.J. 1969. Theoretical and practical problems of Western Yiddish dialectology. In *The Field of Yiddish, Third Collection*, M.I. Herzog, W. Ravid & U. Weinreich (eds.). The Hague: Mouton.

Jakobson, R. & Halle, M. 1964. The term Canaan in medieval Hebrew. In *For Max Weinreich on his Seventieth Birthday*, L.C. Dawidowicz (ed.), 147–172. Berlin: Walter de Gruyter.

Jakobson, R. & Halle, M. 1985. The term Canaan in medieval Hebrew. In *Roman Jakobson: Selected Writings VI: Early Slavic Paths and Crossroads: Part Two: Medieval Slavic Studies*, S. Rudy (ed.), 858–886. Berlin: Mouton de Gruyter. doi:10.1515/9783110863901

Jochnowitz, G. 1978. Judeo-Romance languages. In *Jewish Languages: Theme and Variations*, H. Paper (ed.), 65–74. Cambridge MA: Association for Jewish Studies.

Jochnowitz, G. 1981. Religion and taboo in Lason Akodesh (Judeo-Piedmontese). *International Journal of the Sociology of Language* 30: 107–117.

Jochnowitz, G. 2010. Western Yiddish in Orange County, New York State. *Les Cahiers du CREDYO* 5.

Jochnowitz, G. 2013. Judeo-Italian, Hebrew component. In *Encyclopedia of Hebrew Language and Linguistics*, G. Khan (ed.). Leiden: Brill.

Jochnowitz, G. *Primo Levi, Linguist*. 18. 06. 2012.
<http://www.jochnowitz.net/Essays/PrimoLevi.html>

Jochnowitz, G. The absence of northern features in the Judeo-Italian of Lombardy and Emilia.
<http://www.jochnowitz.net/Essays/JudeoItalian.html>

Kaganovitch, A. 2008. The Bukharan Jewish diaspora at the beginning of the 21st century. In *Bukharan Jews in the 20th Century: History, Experience and Narration*, I. Baldauf, M. Gammer & T. Loy (eds), 111–116. Wiesbaden: Reichert Verlag.

Kakitelashvili, K. 2012. Defining Georgian Jewry: Reconstruction of history through identity building narrative. *Civilizaciuri Ziebani* 42.

Katz, D. 2004. *Words on Fire: The Unfinished Story of Yiddish*. New York NY: Basic Books.

Klier, J. 2010. Pale of settlement. *YIVO Encyclopedia of Jews in Eastern Europe*.

Krivoruchko, J.G. 2011. Judeo-Greek in the era of globalization. *Language and Communication*. 31(2): 119–129. doi:10.1016/j.langcom.2010.08.004

Kushner-Bishop, J. 2004a. From shame to nostalgia: Shifting language ideologies in the Judeo-Spanish maintenance movement. In *Proceedings of the Twelfth British Conference on Judeo-Spanish Studies, 24–26 June, 2001: Sephardic Language, Literature and History*, H.S. Pomeroy & M. Alpert (eds), 23–31. Leiden: Brill.

Kushner-Bishop, J. 2004b. More Than a Language, a Travel Agency: Ideology and Performance in the Israeli Judeo-Spanish Revitalization Movement. PhD dissertation, University of California, Los Angeles.

Kuznitz, C.E. 2004. Yiddish studies. In *Oxford Handbook of Jewish Studies*, M. Goodman (ed.). Oxford: OUP.

Kuznitz, C.E. 2014. *YIVO and the Making of Modern Jewish Culture: Scholarship for the Yiddish Nation*. Cambridge: CUP. doi:10.1017/CBO9781139013604

Labov, W. 2008. Unendangered dialects, endangered people. In *Sustaining Linguistic Diversity: Endangered and Minority Languages and Language Varieties (Georgetown University Round Table on Languages and Linguistics)*, K.A. King, N. Schilling-Estes, L. Fogle, J. L. Lia & B. Soukup (eds), 219–238. Washington DC: Georgetown University Press.

Lepschy, A.L. & Lepschy, G. 1998. *The Italian Language Today*. Chicago IL: New Amsterdam Books.

Levi, P. 1985. *The Periodic Table*. London: Michael Joseph.

Lewis, G. 1999. *The Turkish Language Reform: A Catastrophic Success*. Oxford: OUP.

Lewis, M., Simons, P., Gary, F. & Fennig, C.D. (eds). 2013. *Ethnologue: Languages of the World*. Dallas TX: SIL International.

Litman, S. 2010. Saving another dying Jewish language before it's too late, *Haaretz*.

Margolis, R. 2011. Hip Hop Khasene: A Marriage between Hip hop and Klezmer. *Studies in Religion/Sciences Religieuses* 40(3): 365–380. doi:10.1177/0008429811408214

Margolis, R. 2013. Choosing Yiddish in the classroom: Montreal's national secular schools, 1910–1950. In *Choosing Yiddish: New Frontiers of Language and Culture*, L. Rabinovitch, S. Goren & H.S. Pressman (eds), 123–140. Detroit MI: Wayne State University Press,

May, S. & Hill, R. In press. Maori-medium education: Current issues and challenges. *International Journal of Bilingual Education and Bilingualism* 8(6).

Moskovich, W. & Ben Oren, G. 1981. The Hebrew-Aramaic and Georgian components in the spoken language of Georgian Jews. Paper presented at the *World Congress of Jewish Studies*.

Mufwene, S.S. 2001. *The Ecology of Language Evolution*. Cambridge: CUP. doi:10.1017/CBO9780511612862

Mufwene, S.S. 2005. Language evolution: The population genetics way. In *Gene, Sprachen und ihre Evolution*, G. Hauska (ed.), 30–52. Regensburg: Regensburg University Press.

Naar, D. 2011. Jewish Salonica and the 'Making of the Jerusalem of the Balkans,' 1890–1943. PhD dissertation, Stanford University, Stanford CA.

Parker, S. 1978. Yiddish schools in North America. In *Case Studies in Bilingual Education*, B. Spolsky & R.L. Cooper (eds), 312–330. Rowley MA: Newbury House.

Perlin, R. 2014. Is the Islamic State exterminating the language of Jesus? *Foreign Policy*, 14 August 2014.

Preston, D.R. 2004. Three kinds of sociolinguistics: A psycholinguistic perspective. In *Sociolinguistic Variation: Critical Reflections*, C. Fought (ed.), 140–158. Oxford: OUP.

Sabar, A. 2008. *My Father's Paradise: A Son's Search for His Jewish Past in Kurdish Iraq*. Chapel Hill NC: Algonquin Books.

Sabar, Y. 2003. Aramaic, once an international language, now on the verge of extermination: Are the days of its last vestiges numbered? In *When Languages Collide: Perspectives on Language Conflict, Language Competition, and Language Coexistence*, B.D. Joseph, J. Stefano, N.G. Jacobs & I. Lehiste (eds), 222–235. Columbus OH: Ohio State University Press.

Sadan, T. 2011. Yiddish on the Internet. *Language and Communication* 31(2): 99–106. doi:10.1016/j.langcom.2010.08.002

Sarhon, K.G. 2011. Ladino in Turkey: The situation today as reflected by the Ladino database project. *European Judaism* 44(1): 62–71. doi:10.3167/ej.2011.44.01.08

Shandler, J. 2006. *Adventures in Yiddishland: Postvernacular Language and Culture*. Berkeley CA: University of California Press.

Shapira, D. 2003. Judeo-Persian. *Jewish language research website*, 8 August 2003. <http://www.jewishlanguages.org/judeo-persian.html>

Shapiro, D. 2010. Juhuri. In *Encyclopedia of Jews in the Islamic World*, N.A. Stillman (ed.), 80–81. Leiden: Brill.

Shneer, D. 2004. *Yiddish and the Creation of Soviet Jewish Culture: 1918–1930*. Cambridge: CUP.

Shternshis, A. 2006. *Soviet and Kosher: Jewish Popular Culture in the Soviet Union*, 1923–1939. Bloomington IN: Indiana University Press.

Spicehandler, E. 1968. A descriptive list of Judeo-Persian manuscripts at the Klau library of the Hebrew Union College. *Studies in Bibliography and Booklore* 8: 114–136.

Spolsky, B. 1987. *Maori-English Bilingual Education*. Wellington: New Zealand Department of Education.

Spolsky, B. 2005. Maori lost and regained. In *Languages of New Zealand*, A. Bell, R. Harlow & D. Starks (eds), 67–85. Wellington: Victoria University Press.

Spolsky, B. 2010. Conditions for language revitalization: A comparison of the cases of Hebrew and Maori. In *Endangered Languages: Critical Concepts in Language Studies*, P.K. Austin (ed.), 177–201. London: Routledge.

Spolsky, B. 2014. *The Languages of the Jews: A Sociolinguistic History*. Cambridge: CUP. doi:10.1017/CBO9781107295292

Spolsky, B. 2015. Survival or loss: Lessons from Hebrew and Jewish language varieties. In *Minority Languages: Threatened to Compete*, P. Rewi & R. Higgins (eds).

Spolsky, B. & Shohamy, E. 1999. *The Languages of Israel: Policy, Ideology and Practice*. Clevedon: Multilingual Matters.

Starck, A. 1994. *Westjiddisch: Mündlichkeit und Schriftlichkeit = Le Yiddish occidental: Actes du colloque de Mulhouse*. Aarau: Verlag Sauerländer.

Starck, A. 2007. *Yiddish: Continuity and Change*. Uppsala: Centre for Multiethnic Research, Uppsala University.

Stillman, N.A. (ed.). 1979. *The Jews of Arab Lands: A History and Source Book*. Philadelphia PA: Jewish Publication Society of America.

Stillman, N.A. (ed.). 1991. *The Jews of Arab Lands in Modern Times*. Philadelphia PA: Jewish Publication Society of America.

Uličná, L. & Polakovič, D. 2013. Knaanic glosses from the perspective of Judeo-Czech studies. In *Knaanic language: Structure and historical background: Proceedings of a conference held in Prague on October 25–26, 2012*, O. Bláha, R. Dittman & L. Uličná (eds), 303–317. Prague: Academia.

Weil, S. 2005. Motherland and fatherland as dichotomous diasporas: The case of the Bene Israel. *Presses Universitaires de Rennes, Rennes*, 91–99.

Weinreich, M. 2008. *History of the Yiddish Language*. New Haven CT: Yale University Press.

Weinreich, U. 1949. *College Yiddish: An Introduction to the Yiddish Language and to Jewish Life and Culture*. New York NY: Yiddish Scientific Institute.

Weiser, K.I. 2011. *Jewish People, Yiddish Nation: Noah Prylucki and the Folkists in Poland*. Toronto: University of Toronto Press.

Weiser, K. 2015. Warsaw: The Jewish "capital". In *Warsaw: The Jewish Metropolis: Essays in Honor of the 75th Birthday of Professor Antony Polonsky*, G. Dynner & F. Guesnet (eds), 298–322. Leiden: Brill.

Werman, G. (ed.). 1996. *The Stories of David Bergelson*. Syracuse: Syracuse University Press.

Yeʿor, B. (ed.). 1985. *The Dhimmis: Jews and Christians under Islam*. Rutherford: Fairleigh Dickinsohn University Press.

Yefroikin, S. 1955. Yiddish secular schools in the United States. In *The Jewish People Past and Present*, Jewish Encyclopedia Handbooks (ed.), 144–150.

Index

This peer-reviewed collection brings together the latest research on language endangerment and language rights. It creates a vibrant, interdisciplinary platform for the discussion of the most pertinent and urgent topics central to vitality and equality of languages in today's globalised world. The novelty of the volume lies in the multifaceted view on the variety of dangers that languages face today, such as extinction through dwindling speaker populations and lack of adequate preservation policies or inequality in different social contexts (e.g. access to justice, education and research resources). There are examples of both loss and survival, and discussion of multiple factors that condition these two different outcomes. We pose and answer difficult questions such as whether forced interventions in preventing loss are always warranted or indeed viable. The emerging shared perspective is that of hope to inspire action towards improving the position of different languages and their speakers through research of this kind.

"This collection of papers brings together world-class established scholars and rising stars from a range of language-related disciplines to address key issues in the field of endangered language studies. I recommend it especially for early-career researchers and students who want an overview of the wide range of scholarship in this field."

Julia Sallabank, *SOAS, University of London*

"This volume is a welcome addition to the fast-growing literature on language endangerment. The editors have collected a set of excellent papers that cover the most important areas of the general topic: causes of and responses to language endangerment; language documentation and revitalization; the interconnected issues of ideology, language policy, and language rights; and issues in assessing language vitality."

Sarah Thomason, *University of Michigan*

JOHN BENJAMINS PUBLISHING COMPANY