STUDIES IN
LANGUAGE
COMPANION
SERIES 142

Responses to Language Endangerment

In honor of Mickey Noonan

Edited by
Elena Mihas
Bernard Perley
Gabriel Rei-Doval
Kathleen Wheatley

JOHN BENJAMINS PUBLISHING COMPANY

Responses to Language Endangerment

Studies in Language Companion Series (SLCS)

This series has been established as a companion series to the periodical *Studies in Language*.

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

Editors

Editorial Board

Volume 142

Responses to Language Endangerment. In honor of Mickey Noonan.
New directions in language documentation and language revitalization
Edited by Elena Mihas, Bernard Perley, Gabriel Rei-Doval and Kathleen Wheatley

Responses to Language Endangerment

In honor of Mickey Noonan

New directions in language documentation and language revitalization

Edited by

Elena Mihas
Bernard Perley
Gabriel Rei-Doval
Kathleen Wheatley
University of Wisconsin-Milwaukee

John Benjamins Publishing Company
Amsterdam / Philadelphia

The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

CIP data is available from the Library of Congress.

Studies in Language Companion Series, ISSN 0165-7763 ; v. 142
ISBN 978 90 272 0609 1 (Hb ; alk. paper)
ISBN 978 90 272 7115 0 (Eb)

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

Table of contents

Dedication

This volume is dedicated to the late Michael (Mickey) Noonan (1947–2009), Professor of Linguistics at the University of Wisconsin-Milwaukee, who contributed a significant body of high quality theoretical and descriptive work to functional-typological linguistics. Among his seminal works are a cross-linguistic study of complementation (2007[1985]), the frequently cited grammar of Lango (1991), and numerous papers on grammatical aspects of Tibeto-Burman languages. Michael Noonan was an editor of *Typological Studies in Languages*, co-editor of *Studies in Language* and its Companion Series, and the founding editor of *Himalayan Linguistics*.

His distinguished record of publications and editorships reflects only a fraction of the work he did as a field linguist, tirelessly promoting objectives of language documentation to "record everything!" (in his own words) and provide practical support in language maintenance and revitalization efforts of endangered language communities. Mickey Noonan's life-long commitment was to the Chantyal-speaking community of Nepal. His excellent dictionary of Chantyal (1999) and a book of children's stories were the result of community-based and community-oriented work. The storybook -the first ever published in the Chantyal language-was distributed free to the schools in three Chantyal-speaking villages.

Perhaps Mickey Noonan's most profound impact has been on the people with whom he worked and whom he taught and mentored. After his death, Mickey's academic and human influence was recognized by many friends, colleagues, language consultants, and students whose lives he touched (Moravcsik 2009; Genetti 2009). For many, his work ethic, dedication to 'small' languages and 'small' people, and his humility became a model to follow. His infectious enthusiasm with regard to studying endangered languages in far-flung regions of the world continues to inspire his former master's and doctoral students.

This volume is special in that it was precipitated by Mickey's plan, which he shared with his colleagues in 2008. The volume was envisioned by Mickey Noonan as a necessary conclusion to the conference on language documentation and revitalization to be held in Milwaukee and hosted by the University of Wisconsin-Milwaukee. In accordance with his proposal, the 26th UWM Symposium on Language Documentation and Revitalization was held in 2011 and an impressive body of conference papers was collected. So this publication not only posthumously celebrates Michael Noonan's life as a Linguist and Teacher but also recognizes his initiative to organize the conference and publish the forum's best papers.

The editors believe that this dedication will be an appropriate way of acknowledging Michael Noonan's legacy in language documentation and revitalization, the field of linguistics whose importance he especially recognized (2006). It is also a reminder that as his brainchild, this volume certainly owes its existence to Michael Noonan, although he didn't live long enough to see it through.

References

Genetti, Carol. Mickey Noonan… *Linguaphiles. Live Journal*. February 28, 2009. http://linguaphiles.livejournal.com/4346926.htm. Accessed 09-26-12.

Moravcsik, Edith. 2009. Describe everything! Grammar writer for a grammar-reading audience: In remembrance of Mickey Noonan. *Linguistic Typology* 13: 353–356.

Noonan, Michael P. 1991. *A Grammar of Lango*. Berlin: Mouton de Gruyter.

Noonan, Michael P. 2006. Grammar writing for a grammar-reading audience. *Studies in Language* 30: 351–365.

Noonan, Michael P. 2007 [1985]. Complementation. In Timothy Shopen (ed.), *Language Typology and Syntactic Description* (2nd edn), Vol 2: *Complex Constructions*, 52–150. Cambridge: Cambridge University Press.

Noonan, Micahel P., Ram Bhulanja, Jag Man Chhantyal, & William Pagliuca. 1999. *Chantyal Dictionary and Texts*. Berlin: Mouton de Gruyter.

Acknowledgements

There are many people that we would like to thank for their support of the 26th Linguistics Symposium at the University of Wisconsin-Milwaukee, held in October 2011. The papers in this volume represent a selection of a small portion of the papers that were presented at this conference, and the editors of this volume, who were also on the Organizing Committee of the conference, would not have been able to organize this conference successfully without the help of other members, especially Hamid Ouali, Co-Chair of the Conference, Edith Moravcsik, Bozena Tieszen, and Sally Noonan. There were also many other faculty members and students, who are too numerous to name here, that also assisted in the details of making the conference such a success, and their efforts were greatly appreciated.

The Conference was made possible with the generous support of many units at UWM, including the Provost's Office, the Graduate School, the College of Letters and Science, the Center for Latin American and Caribbean Studies, the Center for International Education, the Electa Quinney Institute, the American Indian Studies Program, the Department of Linguistics, the Department of Anthropology, the Department of Spanish and Portuguese, the Department of French, Italian, and Comparative Literature, the Sam and Helen Stahl Center for Jewish Studies, the Center for Celtic Studies, the Center for 21st Century Studies, the UWM Bookstore, and the Student Union Event Planning Services.

We also want to express our gratitude to the external reviewers who dedicated their time and expertise to reviewing the papers and responding with useful and constructive feedback to the authors: Ewa Czaykowska-Higgins at the University of Victoria, Lise Dobrin at the University of Virginia, Pamela Downing at the University of Wisconsin-Milwaukee, Lenore Grenoble at the University of Chicago, Paul Krosrity at the University of California-Los Angeles, Mary Linn at the University of Oklahoma, Aida Martinovic-Zic at Montgomery College, Barbra Meek at the University of Michigan, Edith Moravcsik at the University of Wisconsin-Milwaukee, Fernando Ramallo at the University of Vigo, Victoria Rau at the National Chung Cheng University in Taiwan, Keren Rice at the University of Toronto, Xoán Paulo Rodríguez Yáñez at the University of Vigo, and Lindsay

Whaley at Dartmouth College. In addition, we would like to thank Cornelis H.J. Vaes, Werner Abraham, and Elly van Gelderen, editors at John Benjamins that supported this project and provided us with valuable feedback.

Finally, we would like to extend our thanks to all of the participants and contributors to the conference, whose enthusiasm and commitment to the revitalization of endangered languages was truly inspiring.

Elena Mihas, Bernard Perley, Gabriel Rei-Doval, & Kathleen Wheatley

Introduction

As the title makes clear, this volume focuses on language endangerment issues, with language endangerment defined as "en mass, often radical shift away from unique, local languages and language practices" (Woodbury 2011: 160). Research on language loss and shift has been a preoccupation of scholars since at least the end of the Second World War, and continued with the subsequent processes of nation building and reshaping in the 1940s and 1950s, as well as the connected processes of development of third world nations and communities. Therefore, the development of disciplines such as language planning or sociolinguistics of society has been crucial to better understand and address the weakening or loss of indigenous and minority languages around the globe.

However, at the beginning of the 1990s preoccupation with language loss made it a compelling task to be urgently attended to by scholars not included in the previously mentioned paradigms, and the field of language revitalization saw a very significant expansion. As a result of this increased interest, questions regarding the best responses to language endangerment have been raised by linguists, sociologists, social psychologists, and native speakers of indigenous languages (e.g. Fishman 1991; Grenoble & Whaley 1998; Hinton & Hale 2001; Lopez 1998; Reyhner et al. 1999; Reyhner et al. 2003). Since then, a special subfield of linguistics that deals with language endangerment issues, documentary linguistics, has gone through a period of dynamic growth. In practical terms, various action items concerning language revival work have been revisited with new eyes and perspectives in recent years (e.g. Austin & Sallabank 2011; Flores Farfán & Ramallo 2010; Grenoble & Furbee 2010; Grenoble & Whaley 2006; Harrison, Rood & Dwyer 2008), aiming to also contribute to the goal of preventing, stabilizing, or reversing the rapidly accelerating local language shift to languages of national or regional stature. Taking stock of the most recent proposals, the key responses to language endangerment have involved (i) language documentation, including creation of a rich multi-genre corpus of recorded connected discourse of the endangered language community and digital archiving of collected data, (ii) establishment of funding bodies to finance documentation projects and, to a lesser extent, language revitalization work, (iii) creation of language education programs by government bodies and other interested parties, and (iv) training of documentary linguists and language teachers, the latter two being roughly subsumed under the domain of language revitalization (Austin & Sallabank 2011).

Language description, equated with the end products of documentation, typically a grammar and a dictionary, is regarded as a supplementary goal of language documentation (Himmelmann 1998).

The current volume further complicates and advances the contemporary perspective, as reflected in its subtitle *New Directions in Language Documentation and Language Revitalization* and evidenced in the content of the volume. In particular, Part I, *Language Endangerment: Challenges and Responses*, offers a general discussion of some of the complex and pressing issues, such as the assessment of the degree of language endangerment (Simons & Lewis, Chapter 1 'The world's languages in crisis: A 20-year update'), the contribution of linguistic scholarship to language revitalization programs (Mithun, Chapter 2 'What can revitalization work teach us about documentation?), the creation of successful language reclamation programs with regard to 'emergent' languages which arise as a result of revitalization efforts after the interrupted transmission (Grenoble, Chapter 3 'Unanswered questions in language documentation and revitalization: New directions for research and 'action'), the training of field linguists and language educators (Genetti & Siemens, Chapter 4 'Training as empowering social action: An ethical response to language endangerment'), and the ethics of fieldwork (Thomason, Chapter 5 'How to avoid pitfalls in documenting endangered languages').

The volume's other section, Part II, *Case Studies in Documentation and Revitalization of Endangered Languages and Languages in Contact*, consists of detailed accounts of fieldworkers and language activists grappling with issues of language documentation and revitalization in the concrete physical and socio-cultural settings of the native-speaker communities. The subtitle of the volume *New Directions in Language Documentation and Revitalization* is also indicative of the imperative for the direct involvement of the language community in the ongoing discussion of the actions undertaken in response to language endangerment.

The volume contains two theory-oriented contributions, from Hildebrandt, Chapter 6 'Converb and aspect-marking polysemy in Nar', and from Jany, Chapter 7 'Grammatical relations in Mixe and Chimariko: Differences and similarities'. Both works highlight the need to document sophisticated grammatical patterns whose linguistic description ultimately enriches language theory, as is the case with the aspect-marking converbial constructions in Nar (Tamangic, Tibeto-Burman), spoken in Nepal, and the expression of grammatical relations in Chuxnabán Mixe (Mixe-Zoquean), the at-risk language of Mexico, and Chimariko, the extinct language of Northern California, USA.

One of the focal issues in documentation of endangered languages is the ethics of data collection and the establishment of partnership relations between the linguist and community members, regarded to be central to the success of documentation projects (Austin & Sallabank 2011: 13). Nakayama & Ono's 'Having a *shinshii*/

shishii[1] 'master' around makes you speak Japanese!: Inadvertent contextualization in gathering Ikema data' (Chapter 8), and Viñas-de-Puig's 'Internal and external calls to immigrant language promotion: Evaluating the research approach in two cases of community-engaged linguistic research in Eastern North Carolina' (Chapter 9) underscore the pivotal importance of forging collaborative relationships with members of the speaker community. The continuous negotiation of the linguist and community researchers' mismatching agendas and identities is illustrated by the fieldworkers' engagements with the native community of Ikema, a dialect of Miyako Ryukyan (Japonic) from Japan, and the immigrant communities of two indigenous languages of Mexico, Tzotzil (Mayan) and Hñähñu (Oto-Manguean), currently spoken in North Carolina, USA.

Of significant importance to both linguists and language activists are documentary and descriptive studies of local languages with small and middle-range numbers of speakers, which are showing signs of convergence onto the dominant language in situations of long-standing language contact and stable bilingualism. In Chapter 10 'Code-switching in an Erzya–Russian bilingual variety: An "endangered" transitory phase in a contact situation', Janurik documents striking structural changes in the versions of Erzya (Finno-Ugric) spoken by 'neo' speakers in the Russian Federation.

Another critical area of research is studies of 'linguistic ecologies' of endangered languages, i.e. socio-cultural and economic settings of native communities, in view of their direct relevance to language shift (Austin & Sallabank 2001: 21). It is commonly observed that local languages fade away when a radical disruption of native speakers' traditional lifeways takes place, by way of 'dislocating' speakers geographically (e.g. due to territorial expropriation of tribal land), economically (e.g. by forcing speakers to switch from hunting and fishing to farming), and/or culturally (e.g. when the homogenizing national ideology places emphasis on being like Us, not Others). In Chapter 11 'Colonialism, nationalism and language vitality in Azerbaijan', Clifton discusses linguistic ecologies and chances of survival of the indigenous languages of Azerbaijan, Talysh and Tat (Iranian), and three North Caucasian (Shahdagh) languages, Budukh, Kryz, and Khinalug.

Training of field linguists and language teachers, and design and implementation of effective teaching methods are a sorely wanting field of study. It is still not clear what teaching models work best or what assessment criteria should be used to determine the success of a revitalization program. It is important to determine how to make training more accessible and more responsive to the needs of field linguists, on the one hand, whose main task is to document the language in situated uses, and the needs of language practitioners, on the other hand, who run language classes in the native communities. The volume's contributions by Genetti & Siemens, Chapter 4 'Training as an empowering social action: An

ethical response to language endangerment' and by Jensen, Jacob & Underriner, Chapter 12 'Revitalizing languages through place-based language curriculum: Identity through learning' tackle these issues. In particular, Genetti & Siemens provide the specifics of the training programs conducted within the participatory community-based research framework by Infield (Institute on Field Linguistics and Language Documentation), University of Santa Barbara, in 2008. Jensen, Jacob, & Underriner elucidate the pedagogy of a place-based curriculum, with its "focus on topics directly related to culture and culture revitalization", developed at the Northwest Indian Language Institute at the University of Oregon and implemented in the indigenous communities of the US Pacific Northwest.

This volume's special concern is with language revitalization and reclamation of autochthonous languages. Notwithstanding the great advances in the theory and practices of language documentation, which have made fieldworkers' engagement with endangered language communities more effective, language revitalization issues remain 'under-theorized' and 'under-researched' (Austin & Sallabank 2011: 22). The present collection explores some of the available solutions. Among the tasks demanding especially urgent attention are production of language materials which 'reconcile full complexity and user-friendliness of grammatical patterns' and 'ways in which languages are special' (Mithun, Chapter 2 'What can revitalization work teach us about documentation?'). Mithun details extant Mohawk (Iroquoian) structures of fascinating complexity which may be lost by the younger generations of bilingual 'neo' speakers (who speak a somewhat simplified version of the traditional language), unless these structures are thoroughly documented in situated contexts and included in the reference grammar materials for community use in the US and Canada. This situation is reported to be difficult to resolve, when the linguist aims to write a comprehensive grammatical description rather than a simpler, pedagogically-oriented account of grammar basics, requested by native language teachers and activists (Austin & Sallabank 2011: 17). In Chapter 5 'How to avoid pitfalls in documenting endangered languages', Thomason gives practical advice on preparation of dictionary materials for the community, based on the author's experience with speakers of Pend d'Oreille (Montana Salish), the indigenous language of Montana, USA. In Chapter 13, Perley's 'Remembering ancestral voices: Emergent vitalities and the future of indigenous languages' offers an evaluative view of the adequacy of responses to language endangerment from the perspective of a linguist/native speaker. Perley outlines concrete steps in the creation of various multi-media projects for indigenous communities. Using the Maliseet (Algonquian) native community as an example, this chapter points in the new direction of revitalization work which includes production of highly entertaining, artfully illustrated 'graphic novels', narrating native stories in Maliseet and English, a series of inspiring native-language stories for television, and television shows with native language voice-overs, accompanied by English subtitles. Graphic

novels are envisioned to be eventually transformed into a tablet device application. Crucially, such multi-media products aim to reflect the modern conventions of Maliseet conversational genre and the use of the native language in contemporary language domains and media.

The ultimate goal of this volume is to offer a forum for academics and members of native speaker communities to take stock of thorny issues and examine the outcomes of the most commonly cited 'responses' to language endangerment, i.e. language documentation, language revitalization, and training. Produced by a variety of authors, including veteran linguists, beginning scholars, and language activists, the contributions to the current volume reflect multiple perspectives and experiences in the field. It is hoped that the current collection will meaningfully contribute to the ongoing conversation on the complexities and practical ways of counteracting language endangerment.

References

Austin, Peter & Sallabank, Julia. 2011. Introduction. In *The Cambridge Handbook of Endangered Languages*, Peter Austin & Julia Sallabank (eds), 1–24. Cambridge: CUP.

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

Flores Farfán, José Antonio & Ramallo, Fernando (eds). 2010. *New Perspectives on Endangered Languages*. Amsterdam: John Benjamins.

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

Grenoble, Lenora & Whaley, Lindsay (eds). 1998. *Endangered Languages: Language Loss and Community Response*. Cambridge: CUP.

Grenoble, Lenora & Furbee, N. Louanna (eds). 2010. *Language Documentation: Practice and Values.* Amsterdam: John Benjamins.

Hale, Ken, Krauss, Michael, Watahomigie, Lucille J., Yamamoto, Akira Y., Craig, Colette, Masayesva, Jeanne & England, Nora. 1992. Endangered languages. *Language* 68:1–42.

Harrison, K. David, Rood, David & Dwyer, Adrienne (eds). 2008. *Lessons from Documented Endangered Languages* [Typological Studies in Language 78]. Amsterdam: John Benjamins.

Himmelmann, Nikolaus. 1998. Documentary and descriptive linguistics. *Linguistics* 36:161–195.

Hinton, Leanne & Hale, Ken (eds). 2001. *The Green Book of Language Revitalization in Practice.* San Diego CA: Academic Press.

Lopez, Daniel. 1998. The parents have to do their part: A Tohono O'odham language autobiography. *International Journal of the Sociology of Language* 132:43–45.

Reyhner, Jon, Trujillo, Octaviana V., Carrasco, Roberto Luis & Lockard, Louise (eds). 2003. *Nurturing Native Languages*. Flagstaff AZ: Northern Arizona University.

Reyhner, Jon, Cantoni, Gina, St. Clair, Robert N. & Parsons Yazzie, Evangeline (eds). 1999. *Revitalizing Indigenous Languages*. Flagstaff AZ: Northern Arizona University.

Woodbury, Anthony C. 2011. Language documentation. In *The Cambridge Handbook of Endangered Languages*, Peter Austin & Julia Salabank (eds.), 159–186. Cambridge: CUP.

PART I

Language endangerment: Challenges and responses

The world's languages in crisis

A 20-year update

Gary F. Simons & M. Paul Lewis
SIL International

"The world's languages in crisis" (Krauss 1992), the great linguistic call to arms in the face of the looming language endangerment crisis, was first delivered in an Endangered Languages Symposium at the 1991 annual meeting of the Linguistic Society of America. Using the best available sources, he surveyed the global situation and estimated that only 10% of languages seem safe in the long term, up to 50% may already be moribund, and the remainder are in danger of becoming moribund by the end of this century. Twenty years later, better information is available. In this paper we use information from the latest edition of the *Ethnologue* (Lewis, Simons & Fennig 2013) to offer an update to the global statistics on language viability. Specifically the data for this study come from our work to estimate the level of every language on earth on the EGIDS or Expanded Graded Intergenerational Disruption Scale (Lewis & Simons 2010). Our finding is that at one extreme more than 75% of the languages that were in use in 1950 are now extinct or moribund in Australia, Canada, and the United States, but at the other extreme less than 10% of languages are extinct or moribund in sub-Saharan Africa. Overall we find that 19% of the world's living languages are no longer being learned by children. We hypothesize that these radically different language endangerment outcomes in different parts of the world are explained by Mufwene's (2002) observations concerning the effects of settlement colonization versus exploitation colonization on language ecologies. We also speculate that urbanization may have effects like settlement colonization and may thus pose the next great threat to minority languages.

1. Introduction

In 1991, Michael Krauss and others participated in a symposium on endangered languages at the annual meeting of the Linguistic Society of America. The compilation of the presentations at that symposium was published a year later in the journal of the society and constituted a call to arms for the linguistics community in the face of the looming language endangerment crisis. Krauss (1992) has been the most cited of those who participated in the symposium and the striking

warning regarding the potential demise of 90% of the world's extant languages has been referred to repeatedly. Using the statistics provided by the 11th edition of the *Ethnologue* (Grimes 1988), which Krauss identified as "by far the best single source available", along with corroborating "guesses" of others with whom he consulted, Krauss estimated that only 10% of the world's languages were safe for the longer term, that 50% might at that time be already moribund, and that the remainder might also become moribund by the end of the 21st century.

Since that time, linguists, anthropologists, language activists, and speaker communities themselves have become increasingly focused on the issue of language endangerment. As the organizers of a recent conference on language endangerment, FEL XV, in Quito, Ecuador, have observed, "Language endangerment is now accepted as an important issue of our times…" (Haboud & Ostler 2011: vi). Numerous publications on the topic have been produced and awareness of the potential for the catastrophic loss of linguistic diversity has reached new heights, sparking considerable interest not only among scholars and practitioners but among the broader public as well. Notable among these is the work of Harmon and Loh (2010) who have built on methods used in ecology for quantifying biodiversity to develop an Index of Linguistic Diversity. Using time-series population data from a sample of 1,500 languages worldwide, they have found that global linguistic diversity has declined 20% over the period 1970–2005.

Krauss noted in his LSA presentation that "statistics on language viability are very hard to come by" (Krauss 1992: 4) and in many respects that continues to be the case 20 years later. In the intervening years, *Ethnologue* has continued to collect and publish data on language vitality, much of which is dated and somewhat idiosyncratic in nature. In the 16th edition of *Ethnologue* (Lewis 2009), serious efforts were made to adjust the categorization scheme used in order to recognize the advent of language revitalization efforts by including a new vitality category, "Dormant", in addition to the previously used Active, Nearly Extinct, Second Language Only, and Extinct labels. This was a small step towards being able to report more accurately the state of vitality of the languages of the world. Nevertheless, the statistical profile of language vitality remained difficult to specify with any certainty because of the reporting delays inherent in the research and data gathering processes, but more significantly because of the lack of a feasible common metric with sufficient precision and granularity by which to assess vitality and endangerment (see for example, Lewis 2006, 2008).

In the latest edition of *Ethnologue* (Lewis, Simons & Fennig 2013), we make significant strides in addressing the lack of statistics on language vitality by, for the first time, providing an estimate of relative safety versus endangerment for every language on earth. This advance is made possible by the introduction and large-scale implementation of the Expanded Graded Intergenerational Disruption Scale (EGIDS) (Lewis & Simons 2010).

2. Methodology

2.1 Expanded Graded Intergenerational Disruption Scale (EGIDS)

The EGIDS builds on the Graded Intergenerational Disruption Scale (GIDS), an 8-level scale that Fishman (1991) developed in order to describe and explain stages in reversing language shift when efforts are made to turn threatened languages into safe ones. The GIDS is well elaborated on the safe end of the scale but has only two levels on the endangered end. By contrast, the Language Vitality and Endangerment (LVE) scale developed by the UNESCO Experts Meeting on Safeguarding Endangered Languages (Brenzinger et al. 2003) identifies four levels of endangerment, but does not distinguish different levels on the safe end of the scale. We have developed the EGIDS by harmonizing the GIDS, the LVE, and the categorization scheme that was being used in *Ethnologue* to form a 13-level scale which recognizes more comprehensively different degrees of vitality over the entire range of the vitality-endangerment continuum.

The basic premise of GIDS is that language shift (ending in language death) happens as a language loses functions in society. To reverse language shift, the community must work to bring those functions back. To guard against future shift, the community can work to add new functions that further strengthen the position of the language. The bulk of Fishman's book consists of case studies describing situations in which this has happened. The magnitude of the numbers in the scale notwithstanding, it has been conventional to view the strongest languages (those with the least disruption and thus the lowest numbers) as being at the top of the scale and the weakest languages (those with the highest levels of intergenerational disruption and the highest numbers) as being at the bottom (e.g. Fishman 2001:466). Thus the basic premise of GIDS can be visually summarized as shown in Figure 1.

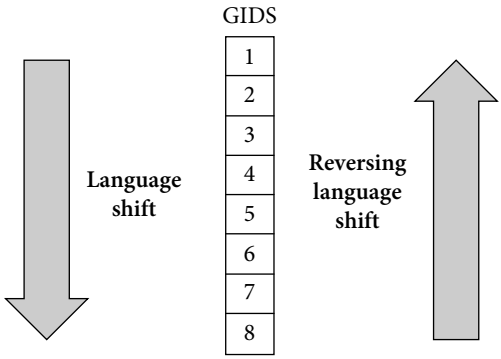

Figure 1. The basic premise of GIDS (Fishman 1991)

The *Ethnologue* (Lewis, Simons & Fennig 2013) is a comprehensive catalog of all known living and recently extinct languages of the world. It gives a basic description of the location and situation of every language listed. In planning for the 17th edition we wanted to provide an estimate for each language as to where it stands on the GIDS scale. When planning for this, however, we encountered the following issues:

1. In order to have a level for every language, we needed to add extinct languages at the bottom of the scale, and in so doing we wanted to keep the *Ethnologue* distinction between dormant languages (which have no fluent speakers but still have an identificational function within an ethnic community) and truly extinct languages (which have no function within any living ethnic community).
2. We observed that in this age of globalization, even official national languages are beginning to feel threatened by the languages of globalization; note, for example, the response of language planners in Sweden in the face of increasing widespread English use (Hult 2005). We have thus added a new level for international languages at the top of the scale.
3. Language endangerment is a huge issue in the world today, but GIDS distinguishes only two levels of endangerment: level 7 in which there is active use of the language but only among adults and level 8 in which the only remaining speakers are "socially isolated old folks" (Fishman 1991: 88). We felt that the users of *Ethnologue* would be better served by a scale that harmonized with the four levels of endangerment recognized in the UNESCO *Atlas of the World's Languages in Danger* (Moseley 2010) which is largely based on UNESCO's LVE assessment framework (Brenzinger et al. 2003).
4. We wanted to add names for the levels, rather than referring to them only by number.

The result is a 13-level scale that we have dubbed EGIDS, for Expanded GIDS (Lewis & Simons 2010). Figure 2 shows the entire scale with a definition for each level. It should be noted that while the scale shown in Figure 2 is congruent with the originally published version, it is not identical; we have been refining the names and definitions of the levels in response to feedback received from users of the scale. In particular, the labels for levels 2 and 3 have been changed in order to align better with the terminology for those language functions identified by William Stewart (1968). The final column of the table gives the corresponding category from the UNESCO language vitality and endangerment (LVE) scale (Brenzinger et al. 2003).

Level	Label	Description	UNESCO
0	International	The language is widely used between nations in trade, knowledge exchange, and international policy.	Safe
1	National	The language is used in education, work, mass media, and government at the nationwide level.	Safe
2	Provincial	The language is used in education, work, mass media, and government within official administrative subdivisions of a nation.	Safe
3	Wider Communication	The language is widely used in work and mass media without official status to transcend language differences across a region.	Safe
4	Educational	The language is in vigorous oral use and this is reinforced by sustainable transmission of literacy in the language in formal education.	Safe
5	Developing	The language is vigorous and is being used in written form in parts of the community though literacy is not yet sustainable.	Safe
6a	Vigorous	The language is used orally by all generations and the situation is sustainable.	Safe
6b	Threatened	The language is still used orally within all generations but there is a significant threat to sustainability because at least one of the conditions for sustainable oral use is lacking.	Vulnerable
7	Shifting	The child-bearing generation can use the language among themselves but they do not normally transmit it to their children.	Definitely Endangered
8a	Moribund	The only remaining active speakers of the language are members of the grandparent generation.	Severely Endangered
8b	Nearly Extinct	The only remaining speakers of the language are elderly and have little opportunity to use the language.	Critically Endangered
9	Dormant	There are no fully proficient speakers, but some symbolic use remains as a reminder of heritage identity for an ethnic community.	Extinct
10	Extinct	No one retains a sense of ethnic identity associated with the language, even for symbolic purposes.	Extinct

Figure 2. Expanded Graded Intergenerational Disruption Scale (Lewis & Simons 2010)

We have retained Fishman's numbering for levels that have an equivalent in GIDS and have used *a* and *b* modifiers to indicate where we have split his levels. We have resisted the temptation to simplify by renumbering the levels from 1 to 13. We feel that it is important to preserve the underlying numbering

scheme of the GIDS for the sake of compatibility with twenty years of prior scholarship and for the face validity that is inherent in following an established standard.

2.2 Generating an EGIDS estimate for every language

In the process of preparing the 17th edition of *Ethnologue* for publication, we have come up with an EGIDS estimate for each of the 7,480 languages (living and extinct) currently tracked in the database. We began by writing a computer script that automatically generated an initial estimate from information available in the database. When the database contained evidence of both development and endangerment, we gave priority to the indicators of threat or endangerment. For instance, a language with published literature was assigned to level 6b rather than 5 if a significant proportion of children are not learning the language.

When the *Ethnologue* database did not contain any information that could provide an initial EGIDS estimate, we consulted UNESCO's *Atlas of the World's Languages in Danger* (Moseley 2010) to see if the language was identified in that work as being in danger. If it was, we followed the assessment in the *Atlas* to give an initial estimate of the EGIDS level by mapping Threatened to level 6b, Definitely Endangered to level 7, Severely Endangered to level 8a, Critically Endangered to level 8b, and Extinct to level 10. After the above steps, approximately one-third of the languages still remained with no estimate of the EGIDS level. In these cases we assigned EGIDS level 6a (Vigorous Oral Use) as the default. In so doing we were following Fishman's (1991: 92) assertion that "the lion's share" of the world's languages are at GIDS 6.

The next step in our process was to send the initial estimates for review to the network of *Ethnologue* contributors and collaborators around the world. Forty-three correspondents, each of whom helps to monitor one or more countries of the world, were asked to review the proposed EGIDS estimates for their countries of focus and to make corrections based on their local and more detailed knowledge. Reviews were returned for 88% of the languages tracked by *Ethnologue* and the corrections entered into the database.

As the last step in the process, we implemented an automated suite of 34 tests to find all the instances of language descriptions in the database in which one of the data elements was potentially inconsistent with the assigned EGIDS level. This caused us to examine about 2,000 language descriptions more closely with the result that hundreds of EGIDS estimates were revised to make them consistent with the story told in the description as a whole. As a sign of the progress that has been made in the process of reviewing and revising the EGIDS

estimates, there were nearly 3,100 languages initially assigned to level 6a (and most of those by default). By the time the 17th edition was published, that number was down to 2,503. No doubt there are errors that persist, but we are confident that the academy as well as the public at large will help us find and correct them as all of the estimates are now published on the *Ethnologue* website.

A final note on methodology must be made. The unit of reporting in *Ethnologue* is the "language in country". That is, each entry in the *Ethnologue* describes the situation of a given language in a particular country. Thus our estimates of the EGIDS level for a language are on a country by country basis. It is these country-specific estimates that were reviewed and corrected. For the analysis below, we are reporting the EGIDS level for the language as a whole. Our method for this is not to take an average of all countries, but to report the highest level (that is, most safe) for any country. The logic here is that if the EGIDS level of a language is taken as a predictor of its likely longevity, then its longevity will be determined by where it is the strongest.

3. Results

3.1 A comprehensive analysis of the state of the world's languages

The EGIDS estimates in the 17th edition of *Ethnologue* represent the first fully comprehensive quantitative analysis of the state of vitality of the world's languages. While many of these estimates should be considered preliminary, the profiles of language vitality that emerge can provide us, for the first time, with a baseline from which trends and patterns can be traced over time as the use of the EGIDS as a metric of ethnolinguistic vitality continues and is refined. Analyses such as that done by Krauss (1992) were necessarily sketchy and impressionistic because the state of our knowledge at that time, even using "the best source available" was not adequate to the task. We believe that the EGIDS can serve as a tool that is feasible to use on a global scale and that provides a better level of granularity and precision than other options that have been developed to date.

3.2 Global results

We start by looking at the global statistics for the distribution of the world's languages by EGIDS levels. Figure 3 shows a histogram of how the languages are distributed by level.

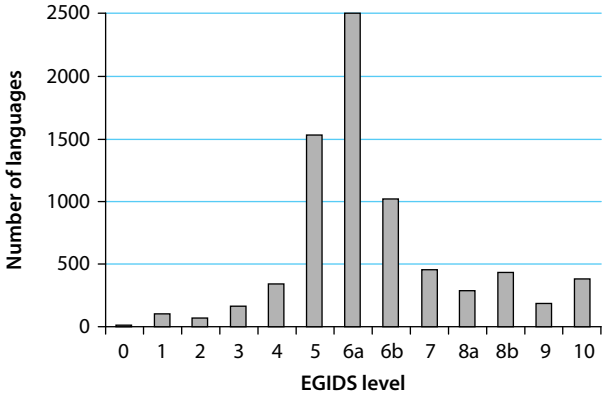

Figure 3. Global distribution of languages by EGIDS level

Table 1 shows the numbers that lie behind the graph, both as counts and percentages. The total number of languages (7,480) represents all the living languages listed in the 2013 update of the ISO 639-3 standard (ISO 2007), plus the languages listed in the standard that have gone extinct since 1950 (which is when the *Ethnologue* began tracking languages).

Table 1. Global distribution of languages by EGIDS level

EGIDS Level	Languages	Percent
0 (International)	6	0.1%
1 (National)	98	1.3%
2 (Provincial)	70	0.9%
3 (Wider communication)	166	2.2%
4 (Educational)	342	4.6%
5 (Developing)	1,534	20.5%
6a (Vigorous)	2,503	33.5%
6b (Threatened)	1,024	13.7%
7 (Shifting)	456	6.1%
8a (Moribund)	286	3.8%
8b (Nearly extinct)	431	5.8%
9 (Dormant)	187	2.5%
10 (Extinct)	377	5.0%
Total	7,480	100.0%

The most striking feature of this distribution is the preponderance of languages at EGIDS level 6a. Globally, 2,503 of the languages of the world are characterized by vigorous oral use. When the count for EGIDS level 6a is combined with the languages at higher, stronger levels (EGIDS 0–5), we see that 4,719 (63%) of the 7,480 languages in use in 1950 are still being passed on to the next generation in a sustainable way. In the discussion which follows, we refer to this group of languages as "vital" languages. In contrast, 1,480 (20%) of the languages of the world are "in trouble" (EGIDS 6b–7). In these languages the norm of complete intergenerational transmission is no longer in effect, but members of the child-bearing generation are still fully proficient in the language so that it would still be possible for a successful revitalization effort to restore intergenerational transmission. Finally, an additional 1,281 (17%) of languages are "dead or dying" (EGIDS 8a–10) since it is too late to restore natural parent-to-child transmission. The restoration of intergenerational transmission would require establishing overt language transmission mechanisms outside the home.

Among the dead and dying languages are 377 (5%) that have been identified as having lost all living speakers and ceasing to serve as a language of identity for an ethnic community (EGIDS 10) in the last six decades. The loss of linguistic diversity represented by the loss of these individual languages is even more staggering if viewed from the perspective of language families. Whalen and Simons (2012) show that with the loss of these languages, we have lost 15% of the linguistic stocks (the largest subgroups of related languages that are reconstructable) that had at least one living member in 1950.

Alarmingly, 2,384 (32%) living languages in the world are currently at some stage in the process of language loss (EGIDS 6b–9). That is more than the number of languages (2,216, 30%) that have experienced enough language development (EGIDS 0–5) to rise above the default stage of vigorous oral use (EGIDS 6a).

3.3 Results by geographical regions

The above global statistics give a sense of the scale of the language endangerment crisis, but they mask the fact that the situation may differ radically from one part of the world to another. To better give a sense of what is happening throughout the world, we present results from our EGIDS survey for each of the 22 geographical regions into which the United Nations divides the world for the purposes of its reporting (United Nations Statistics Division 2011). Each language occurs only once in the regional statistics. Thus when a language is used in multiple regions, we have counted it with the region in which its primary country (as identified in the *Ethnologue*) is located.

Table 2 provides data on the number of languages in each region according to the three summary categories of "Vital" (EGIDS 0–6a), "In Trouble" (EGIDS 6b–7), and "Dead or Dying" (EGIDS 8a–10). The areas are ranked from most to least by the number of dead or dying languages. The top of the table thus shows the regions that have been most heavily impacted by the language endangerment crisis.

Table 2. Geographic regions by number of dead or dying languages (most to least)

Region	Total languages	Vital	In trouble	Dead or dying
Australia and New Zealand	388	35	36	317
South America	521	184	133	204
Northern America	266	18	85	163
South-Eastern Asia	1,277	734	412	131
Melanesia	1,071	811	179	81
Middle Africa	688	555	77	56
Western Africa	899	794	49	56
Central America	332	198	86	48
Southern Asia	680	502	131	47
Eastern Asia	292	152	106	34
Eastern Europe	122	64	25	33
Northern Africa	154	92	29	33
Eastern Africa	390	318	51	21
Western Asia	93	45	32	16
Northern Europe	50	38	2	10
Southern Africa	53	41	3	9
Southern Europe	67	46	13	8
Western Europe	53	39	9	5
Caribbean	24	17	3	4
Micronesia	27	19	5	3
Central Asia	14	9	3	2
Polynesia	19	8	11	0
Total	*7,480*	*4,719*	*1,480*	*1,281*

Heading the list is Australia and New Zealand with 317 dead and dying languages. Next come South America (204), Northern America (163), South-Eastern Asia (131), and Melanesia (81). (Northern America, as distinct from North America, comprises Bermuda, Greenland, Saint Pierre and Miquelon, Canada and the United States; the UN regions group Mexico with Central America.)

With 896 out of the total of 1,281, these five regions account for over two-thirds of the dead and dying languages in the world.

Table 3 presents the same data in a different way. The counts are converted to percentages and the regions are ranked from most to least by the percentage of vital languages. In this listing, Northern America assumes the bottom position with only 7% vital languages. Then come Australia and New Zealand (9%) and South America (35%). These three regions also have the highest percentages of dead and dying languages (61%, 82%, and 39%, respectively).

Topping the list in Table 3 as the part of the world least impacted by language endangerment is sub-Saharan Africa in which the three regions of Western,

Table 3. Geographic regions by percentage of vital languages (most to least)

Region	% Vital	% In trouble	% Dead or dying
Western Africa	88%	5%	6%
Eastern Africa	82%	13%	5%
Middle Africa	81%	11%	8%
Southern Africa	77%	6%	17%
Northern Europe	76%	4%	20%
Melanesia	76%	17%	8%
Southern Asia	74%	19%	7%
Western Europe	74%	17%	9%
Caribbean	71%	13%	17%
Micronesia	70%	19%	11%
Southern Europe	69%	19%	12%
Central Asia	64%	21%	14%
Northern Africa	60%	19%	21%
Central America	60%	26%	14%
South-Eastern Asia	57%	32%	10%
Eastern Europe	52%	20%	27%
Eastern Asia	52%	36%	12%
Western Asia	48%	34%	17%
Polynesia	42%	58%	0%
South America	35%	26%	39%
Australia and New Zealand	9%	9%	82%
Northern America	7%	32%	61%
Global	*63%*	*20%*	*17%*

Eastern and Middle Africa all have more than 80% of their languages in the vital category. Interestingly Melanesia (which ranked fifth in terms of most dead and dying languages) ranks sixth in this list with 76% vital languages, due to the large number of vital languages in Papua New Guinea.

Tables 2 and 3 make it clear that the language endangerment story is very different in different parts of the world. In Australia and the Americas, the crisis has been running its course with devastating consequences, while in sub-Saharan Africa it has yet to hit the radar screen as a crisis. Throughout Asia, Europe, and other regions of the Pacific the situation is between these extremes, but tends much more toward the vital than the dying.

4. Discussion

4.1 Krauss's warning: Is it coming true?

Our findings show that Krauss's estimate in 1992 that 50% of languages were doomed or dying was too dire. With very incomplete data, he sought to estimate the percentage of languages that were no longer being passed down from parents to their children. He noted that "the Grimeses themselves [editors of the *Ethnologue* at that time] might agree that as many as 20% of the world's languages are already moribund. However, two other linguists with wide experience have both independently guessed, along with me, that the total may be more like 50%" (Krauss 1992: 6). Twenty years later we have, for the first time, vitality estimates for all the world's languages. Our finding is that out of 7,103 living languages (EGIDS 0–9), 1,360 (or 19%) are not being learned by children (EGIDS 7–9).

Krauss's predictions were certainly on track in those regions where language shift and loss are most extreme. Working with the data he had, and from his experience largely in Northern America, Krauss's pessimistic predictions are understandable. Indeed, our current data indicate that 78% (207 out of 266) of the languages of Northern America are either already extinct or not being learned by children (EGIDS 7–10), as are 85% (329 of 388) in Australia and New Zealand. Three other regions are approaching the 50% level: South America (48%, 242 of 521), Polynesia (47%, 9 of 19), and Western Asia (41%, 38 of 93).

For the other 16 regions in the world, the proportion of languages that are already extinct or not being learned by children (EGIDS 7–10) ranges from 30% in Eastern Europe (37 of 122) down to 8% in Eastern Africa (31 of 390) and Western Africa (69 of 899). The language ecologies in these other parts of the world are

considerably different from the situations in the Americas and Australia. This in no way diminishes the relevance of Krauss's warning since there are minority languages under threat in all parts of the world. However, on a global scale the threat does not yet reach the level suggested by Krauss. The greater scope and refinement provided by the global EGIDS data gives us a more nuanced understanding and, hopefully, the ability to respond to each situation more strategically and appropriately.

4.2 Mufwene's colonization types: A possible explanation

Extrapolating from what was already evident in Australia and the Americas, Krauss considered it to be "a plausible calculation that – at the rate things are going – the coming century will see either the death or the doom of 90% of mankind's languages" (Krauss 1992: 7). But the global evidence does not seem to be bearing this out. Was it a plausible extrapolation? We believe that the work of Salikoko Mufwene offers an explanation as to why it was not.

Mufwene (2002) has proposed that the outcomes of language contact correspond in great measure to the pattern of colonization which was predominant in that part of the world. He has identified three colonization types: trade, exploitation, and settlement. His proposal, very briefly stated, is that "Each colonisation style has determined particular patterns of interaction between the colonisers and the indigenous populations as well as the particular kind of economic structure that is now in place" (Mufwene 2002: 168). In terms of the dynamics of language maintenance and shift, Mufwene asserts that "European colonial languages have endangered other languages, or driven them to extinction, typically in settlement colonies, not in exploitation nor in trade colonies." (Mufwene 2002: 168).

Mufwene identifies trade contact as the earliest colonization type to emerge. In this pattern of contact, there was occasional language contact as ships periodically landed at ports of call to collect trade goods. Contact languages emerged for conducting business, but contact was not prolonged and did not lead to language shift. In contrast to this, exploitation colonies involved on-going residence by Europeans in plantations or trading centers, but they did not come in large numbers nor did they settle permanently. Language contact was prolonged, but it was not deep. This kind of contact has resulted more often in the maintenance and retention of local languages and the addition of the colonizers' language as a second language in the repertoire of those who were colonized.

In settlement colonies, Europeans came in large numbers, bringing their families to establish a new life in a new land. Language contact was both prolonged

and deep, resulting in profound language shift. Mufwene identifies the marked pattern of language shift that is the after effect of settlement colonization as the predominant explanation of language endangerment:

> Especially noteworthy about settlement colonies is the fact that they gradually produced local or regional monolingualism, favouring the language of the colonising nation but dooming to extinction the languages brought by the Africans … and Europeans originating from countries other than the colonising one …. Native Americans lost their languages either because they were decimated by diseases and wars, or because they were forced to relocate to places where they could not continue to speak their languages, or because they eventually got to function in the new, European-style economic world order which imposed a new language of business and industry. Unlike trade colonies, settlement colonies everywhere gradually evolved to some form of economic (and social) integration that has endangered languages other than those of the colonising European nation, or one adopted by it. (Mufwene 2002: 169)

On general inspection of the results in Tables 2 and 3, Mufwene's proposal seems quite plausible. The places where language loss has been the most profound – Australia, Canada, and the United States – are also places where virtually all of the land was settled by the colonizers, thus displacing the indigenous inhabitants. By contrast, the regions of sub-Saharan Africa and Melanesia, where language loss has been minimal by comparison, were not settled by the colonizers, but were only exploited for the benefit of the home country. Thus it is plausible, and is being argued by some (Bagamba & Boone 2011; Landweer 2012; Landweer & Unseth 2012), that in these regions we would not expect to see the kind of language loss predicted by Krauss, since his prediction is based on an extrapolation of the outcome in regions that were dominated by settlement colonization. A correlation analysis of the colonization patterns that were typical of particular regions or countries with the profile of current EGIDS estimates for the languages in each context could be done to develop concrete evidence that could support Mufwene's hypotheses.

4.3 Urbanization: The next big threat?

Global politics have changed dramatically over the past century with the result that settlement colonization no longer poses much threat of causing new language endangerment in the future. But that is not to say that minority languages are now safe. They are facing a very real threat in this century from a fourth pattern of economic contact with the external world, namely, urbanization. For this reason, linguists should still be giving heed to the warning given by Krauss.

The dynamics of extended contact in urban settings seem similar to those described by Mufwene for the settlement colonization pattern. We would thus

expect similar outcomes. Interestingly, the power dynamics and the process of contact in urbanization is the reverse of what it is for settlement. In settlement colonization, more powerful outsiders moved in and pushed local residents off their land against their will. In urbanization, the less powerful are willingly leaving their ancestral territories and being pulled into urban centers where they are brought into extensive contact not only with the more powerful but with many others who are equally disempowered. The need to acquire proficiency in the dominant languages of the urban centers is posing a new threat to the vitality of minority languages as large numbers of people are moving from rural to urban areas. Given the UN estimate that from 2007 to 2050 the global proportion of urban population will increase from 49% to 70% (United Nations 2008), we can anticipate that the pressures on minority language speakers toward language shift will only increase in the coming decades. It would behoove the linguistics community to give more attention to understanding the mechanisms of language maintenance versus language loss in the context of urbanization.

5. Conclusions

Krauss's "call to arms" in 1992 has helped to mobilize many in the linguistics community to engage in activities aimed at preserving endangered languages and cultures. Krauss's analysis, based on the best evidence available at the time, has been shown to be largely accurate for the parts of the world he and his correspondents were most familiar with (Northern America and Australia), but overall represents an overly pessimistic representation of the state of the languages of the world based on our analysis 20 years later.

The development and global implementation of the Expanded Graded Intergenerational Disruption Scale (EGIDS) enables us to have a much better view of the endangerment situation. While it confirms that language loss predominates in certain regions, the global analysis reveals that there are even more parts of the world in which language maintenance is far more prevalent than language loss.

This analysis has enabled us to confirm that, as Fishman predicted, the largest number of the languages of the world, 63%, are safely maintained in everyday oral use in their communities (EGIDS 6a) or are at a stronger level of development and recognition (EGIDS 0–5). Nevertheless, the statistics also reveal that 32% of the world's languages are in some stage of loss or shift (EGIDS 6b–9) and that 5% of the languages in use in 1950 are now completely extinct (EGIDS 10). This proportion (37%) of languages on the falling end of the scale (EGIDS 6b–10) is more than the 30% that are in some stage of development beyond oral use alone (EGIDS 0–5).

This analysis is preliminary since it is based on our first attempt to estimate the status of every language on earth for inclusion in the latest edition of *Ethnologue*. However, we trust that the results are adequate to begin serving as a baseline for future studies. As additional data on undocumented languages are gathered and as the existing EGIDS estimates are refined and improved, we expect that a much sharper image of the state of the world's languages will emerge. This improved understanding has potential to serve both scholars and members of endangered language communities alike by helping them to better assess the level of disruption of intergenerational transmission.

References

Bagamba, Araali & Boone, Douglas. 2011. Challenges to applying the EGIDS in north-eastern Democratic Republic of Congo. Paper presented at American Association for Applied Linguistics, Annual Meeting. Chicago IL.

Brenzinger, Matthias; Yamamoto, Akira; Aikawa, Noriko; Koundiouba, Dimitri; Minasyan, Anahit; Dwyer, Arienne; Grinevald, Colette; Krauss, Michael; Miyaoka, Osahito; Sakiyama, Osamu; Smeets, Riek & Zepeda, Ofelia. 2003. *Language Vitality and Endangerment.* Paris: UNESCO Ad Hoc Expert Group Meeting on Endangered Languages.

Fishman, Joshua A. 1991. *Reversing Language Shift.* Clevedon: Multilingual Matters.

Fishman, Joshua A. (ed.). 2001. *Can Threatened Languages Be Saved? Reversing Language Shift Revisited: A* 21st *Century Perspective* [Multilingual Matters 116]. Clevedon: Multilingual Matters.

Grimes, Barbara F. (ed.). 1988. *Ethnologue: Languages of the World,* 11th *edn.* Dallas TX: SIL International.

Haboud, Marleen & Ostler, Nicholas (eds). 2011. *Endangered Languages: Voices and Images. Quito: Foundation for Endangered Languages XV Annual International Conference.* Pontificia Universidad Católica del Ecuador.

Harmon, David & Loh, Jonathan 2010. The Index of Linguistic Diversity: A new quantitative measure of trends in the status of the world's languages. *Language Documentation and Conservation* 4: 97–151. ⟨http://scholarspace.manoa.hawaii.edu/bitstream/10125/4474/1/harmonloh.pdf⟩

Hult, Francis M. 2005. A case of prestige and status planning: Swedish and English in Sweden. *Current Issues in Language Planning* 6(1): 73–79.

ISO. 2007. *ISO 639–3:2007: Codes for the Representation of Names of Languages – Part 3: Alpha-3 Code for Comprehensive Coverage of Languages.* ⟨http://www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=39534⟩. Registration authority: ⟨http://www.sil.org/iso639-3/⟩ (15 April, 2012).

Krauss, Michael. 1992. The world's languages in crisis. *Language* 68(1): 4–10.

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

Landweer, M. Lynn & Unseth, Peter. 2012. An introduction to language use in Melanesia. *International Journal of the Sociology of Language* 214: 1–3.

Lewis, M. Paul. 2006. Towards a categorization of endangerment of the world's languages. *SIL Electronic Working Papers* 2006–002. ⟨http://www.sil.org/silewp/abstract.asp?ref=2006–002⟩

Lewis, M. Paul. 2008. Evaluating endangerment: Proposed metadata and implementation. In *Sustaining Linguistic Diversity: Endangered and Minority Languages and Language Varieties,* Kendall King, Natalie Schilling-Estes, Lyn Fogle, Jia Lou & Barbara Soukup (eds), 35–49. Washington DC: Georgetown University Press.

Lewis, M. Paul (ed.). 2009. *Ethnologue: Languages of the World,* 16th *edn.* Dallas TX: SIL International.

Lewis, M. Paul & Simons, Gary F. 2010. Assessing endangerment: Expanding Fishman's GIDS. *Revue Roumaine de Linguistique* 55(2): 103–120. ⟨http://www.lingv.ro/resources/scm_images/RRL-02-2010-Lewis.pdf⟩

Lewis, M. Paul, Simons, Gary F. & Charles D. Fennig (eds). 2013. *Ethnologue: Languages of the World,* 17th *edn.* Dallas TX: SIL International. ⟨http://www.ethnologue.com⟩

Moseley, Christopher. 2010. *Atlas of the World's Languages in Danger,* 3rd *edn.* Paris: UNESCO Publishing. Interactive edition: ⟨http://www.unesco.org/culture/languages-atlas/⟩

Mufwene, Salikoko S. 2002. Colonisation, globalisation, and the future of languages in the twenty-first century. *International Journal on Multicultural Societies* 4(2): 162–193.

Stewart, William A. 1968. A sociolinguistic typology for describing national multilingualism. In *Readings in the Sociology of Language,* Joshua A. Fishman (ed.), 530–553. The Hague: Mouton.

United Nations. 2008. *World Urbanization Prospects: The 2007 Revision.* New York NY. ⟨http://www.un.org/esa/population/publications/wup2007/2007WUP_ExecSum_web.pdf⟩ (20 April 2012).

United Nations Statistics Division. 2011. Composition of macro geographical (continental) regions, geographical sub-regions, and selected economic and other groupings. ⟨http://unstats.un.org/unsd/methods/m49/m49regin.htm⟩ (19 April, 2012).

Whalen, Douglas. H. & Simons, Gary F. 2012. Endangered language families. *Language* 88(1): 155–173.

What can revitalization work teach us about documentation?

Marianne Mithun
University of California, Santa Barbara

As language documentation gains recognition as an important methodology for linguistics, and as communities mount ever more impressive revitalization projects, the interests of academic and community scholars are converging. It is useful to look to those involved in revitalization for their views on what they treasure most in the existing records of their languages and what they wish were there. Decisions about documentation are tightly bound up with ideas about what constitutes the essence of a language. If a language is viewed as encompassing such things as discourse structure, styles of interaction, constructions that meld structure and substance, prefabricated collocations and idiomatic expressions, recurring lexical choices, and conventionalized prosodic structures, then all of these must be part of the record.

1. Introduction

Normal science pushes us to address the questions we are currently asking. It can move knowledge ahead, as new answers raise new questions. But the escalating endangerment of languages all over the world raises important issues for normal work in linguistics. On the one hand, the languages may not be there when new questions develop. On the other, the kinds of information that are most valuable to the communities whose languages they are, and that may become even more valuable to them in the future, may not be the focus of current science. Endangerment brings a special responsibility for scholars, inside and outside of local communities, to think beyond immediate interests. Thoughtful documentation is a more crucial aspect of the field than ever before.

Fortunately, recognition of the importance of good documentation is now widespread, and attention is being directed not only at the technical aspects of the work, but also content. The audiences for language documentation are varied, evolving, and merging. They include both general linguists and those working in more specialized areas. Ever more importantly, they include scholars from within

the communities, with many of the same kinds of interests as those outside, but also with a special stake in the language. There are now impressive revitalization projects underway around the globe, and many of the people involved in them have insight into the kinds of language documentation that will be most valuable to communities as circumstances evolve. They can offer important advice on what kinds of material they have found the most useful for their work, as well as what they wish were in the record but is not.

Michael Noonan (2005) wisely distinguished two main mechanisms leading to language endangerment: *language shift* and *linguistic convergence*. With the first, *shift*, people simply stop speaking the language. For understandable economic and social reasons, they want and need to be able to communicate in the encroaching language and ensure that their children can do the same, so they use that language in more and more situations. Often shift progresses below the level of consciousness of speakers: they themselves are bilingual, but they are sometimes surprised to discover that the next generation does not know the traditional language. In the second scenario, *convergence*, distinctive characteristics of one language are lost by gradual drift toward the grammatical and semantic categories of another. As Noonan remarks, convergence is in a sense more insidious than shift; the language survives, but it loses its distinctiveness. As he notes, convergence is not always obvious from grammatical descriptions, because grammarians are often selective, consciously or unconsciously, when they decide what should go into a grammar, describing only structures deemed "typical" of the language.

Revitalization work is raising our awareness of the kinds of information that should be considered for documentation in both situations: shift and convergence. And as the field of linguistics has evolved, it has become increasingly clear that often the very kinds of documentation valued by communities can contribute significantly to a deeper understanding of the nature of language in general and the forces that shape it.

2. Shift

Language shift can be motivated by a variety of forces, often operating in concert. Bilingualism in an encroaching language makes sense as a way to survive economically and socially in an evolving world. As a consequence, the traditional language can become all but invisible. Children may rarely see it or hear it around them. Both contributing to this situation and as a result of it, the language may be undervalued by community members, by people outside of the community, or both.

2.1 Counteracting invisibility

To counteract loss through invisibility, an obvious strategy is to increase the presence of the traditional language. Certain kinds of documentation can contribute to this effort. Mohawk communities in Quebec, Ontario, and New York State are providing some wonderful examples of what can be done.

One tactic is to increase its visual presence. In Kahnawà:ke, Quebec, for example, intersections are marked with bilingual stop signs: the familiar red hexagonal signs say both *Stop* and *Tésta'n*. Over the door of the church is a sign *Ononhsatokénhti* ('it is house holy'). At the entrance to the bank is a sign whose top line reads *Tsi iehwistaientáhkhkwa'* ('place one lays down money with'), above *Caisse poplulaire Kahnawake*. Along a road is a billboard *Sheia'tánerenk tóka' shenorónhkhwa'* 'Buckle them up if you love them', with cute drawings of babies and children buckled into car seats. There are plaques around town commemorating the *Ratirista'kehró:non'* the heroic ironworkers. For those engaged in language documentation, it can be useful to think of the kinds of vocabulary that might serve as useful community resources for heightening the visibility of the language, whether or not such terms seem to be of current theoretical significance.

Another strategy for heightening awareness of the language is to increase its aural presence. Many factors enter into the language choices bilinguals make in particular settings and with particular people. Speakers can decide to use the language as often as possible. It may not be easy, but it need not require special funds. There is nothing like being surrounded by a language to make a person want to know it. The language immediately becomes more relevant. Even those who are not first-language speakers can make a difference. In a community where residents have dogs and talk to them, for example, speakers and learners alike can canvass the community to collect traditional pet names. The Mohawk dog names *Shentáhsa* 'the tail one, Tail' (for male dogs) and *Skentáhsa* 'Tail' (for female dogs) even show some grammatical structure, with different gender prefixes. Other traditional dog names show similar structure: *Shahónhta* 'Ears', *Shanén:ia* 'Rocky', *Shanénhsta* 'Corny'. Dog owners can learn simple commands and fill the air with them, such as Mohawk *Sátien!* 'Sit!', *Sá:rat!* 'Lie down!', *Satkarhátho!* 'Roll over!', *Tatsé:na!* 'Catch!', and *Tesatkarhaté:ni!* 'Turn around!' They can praise their dogs if so inclined: *Senahskwí:io!* 'Good dog!' (literally 'You are a good animal'). Learning a repertoire of such phrases is not too hard, and dogs are much more tolerant of a learner's pronunciation than many people. Documentation of such language need not take long, but it can yield observable results.

Another useful project is to collect phrases that everyone can use in daily interaction, like Mohawk *Kwé:* 'Hi', *Iawékon* 'It's delicious', and *Niá:wen* 'Thank you'. Assembling lists of expressions used most often on a daily basis can provide

another fruitful research opportunity for community youth and others. What is said in the traditional language can be learned and used as is. Identifying what is said most often in the encroaching language is important as well, since it reflects the things people will probably want to say. Counterparts for those expressions can then be sought in the traditional language.

Other kinds of language that even non-speakers can learn readily, and that can serve as a marker of identity and respect for the language, are interjections. One common Mohawk expression is *Háo'ki* 'Come on'.

(1) Interjections
Háo'ki tetewatskà:hon.
'Come on let's eat.'

Like most languages, Mohawk is rich in such small expressions: *Hátskwi, Hánio, Ótsta', Wáts, Wá'tsik, Á:ke, Akí:*, *Atió:*, and many more. Anyone listening carefully to Mohawk conversation will notice them everywhere. Some are easily translated. *Atió:* is said when someone feels something cold: 'Brrrr'. *Akí:* signals pain: 'Ouch'. Others have meanings like 'Gee whiz' or 'Wait a minute'. Many children who have grown up in Mohawk-speaking communities without learning Mohawk as a first language still use such expressions appropriately.

But expressions like these often did not find their way into the field notes of earlier linguists. These gaps are usually the first thing mentioned by those working to revitalize languages. Three very active individuals spearheading revitalization projects in North America, Daryl Baldwin (Myammia (Illinois) project in Oklahoma), Megan Lukaniec (Wendat (Huron) project in Quebec), and Richard Zane Smith (Wyandot project in Oklahoma) have each cited everyday expressions as the most valuable information they could have, information that is represented only scantily, if at all, in the records of the language they have to work with.

2.2 Empowering more serious learners

Ambitious learners involved in revitalization programs often comment that they know a lot of words, but not how to talk. For those eager for more than short expressions, systematic documentation of basic patterns can serve as the foundation for learning to understand and say new things, to use the language creatively. Yes/no questions, for example, tend to be relatively regular cross-linguistically. In Mohawk, such questions are formed simply with the particle *ken*.

(2) Yes/no questions

Iawékon.	'It's delicious.'
Iawékon ken?	'Is it delicious?'
Iokennó:ron.	'It's raining.'
Iokennó:ron ken?	'Is it raining?'

Learners can get considerable mileage out of this one construction. But to continue the conversation, they will eventually want to know how to answer in the negative. There are clear patterns here as well.

(3) Negation

Iawékon.	'It's delicious.'
Iáh, iáh teiawékon.	'No, it doesn't taste good.'
Iokennó:ron.	'It's raining.'
Iah, iáh teiokennó:ron.	'No, it's not raining.'

The range of sentences learners can use will be increased greatly if they can talk about different persons.

(4) Persons

Katonhkária'ks	'I'm hungry'
Satonhkária'ks	'You're hungry'
Satonhkária'ks ken?	'Are you hungry?'
Kenòn:we's	'I like it'
Senòn:we's	'You like it'
Senòn:we's ken?	'Do you like it?'

In some linguistic traditions, it has been customary to provide no more examples than are necessary to demonstrate the point under discussion. Those documenting the language might assume that the examples of polarity questions in (2), negation in (3), and first and second persons in (4) would be sufficient. But teachers and learners alike need more in order to see the patterns for themselves and absorb them. And there may be limits to the patterns that are not obvious from a single example. The Mohawk negative, for example, is signaled by a different prefix if other prefixes are present. Also, as in many languages, the forms of individual morphemes, like the pronominal prefixes in (4), have different shapes in different contexts.

Documentation of robust patterns like these is fundamental for revitalization. But word lists and basic grammar alone do not make full speakers. Much of knowing a language is knowing what to say in a given situation. Language comes in recurring, situated chunks. Piece-by-piece translations of sentences from the encroaching language may be technically grammatical but completely unidiomatic, simply not what speakers say. In each of the examples below, the first two lines provide idiomatic English and Mohawk counterparts. The third lines provide literal translations of the idiomatic Mohawk. The idiomatic Mohawk could never be derived from the idiomatic English.

(5)

English:	'So how's everything?'			
Mohawk:	*Kwé:.*	*Hátskwi*	*oh*	*ní:ioht?*
Literally	Hi.	so then	how	so it is

(6) English: 'Watch out!'
Mohawk: *Se'nikòn:rarak !*
Literally 'Keep mind setting on it!'

(7) English: 'I agree'
Mohawk: *Né:* *ki'* *wáhi'*.
Literally that.aforementioned in.fact TAG

People in all cultures learn social formulas. There are various ways in Mohawk to invite someone to take a seat next to you, for example. (All of these examples were drawn from spontaneous Mohawk conversation among skilled first-language speakers. Each line represents a separate intonation unit or prosodic phrase.)

(8) 'Sit here!'

a. *Kén:* *satien!*
here you set yourself

b. *Kén:* *sanitskó:ten!*
here you bum stand

There are certain ways of extending invitations.

(9) 'Come join us!'
Ká:ts, *takwatia'tárhahs*
come put yourself bodily in for us.'

(10) 'What do you say we go play some golf?'
Ó:nen *kati'* *ken*
now then Q
tentsitewahthennoʔókha'?
there will back you all and I go to ball hit

There are ways of requesting permission.

(11) 'May I have a cup of coffee?'
Enwá:ton' *ken* *enkatathnekárhahse'?*
it will be possible Q I will myself liquid serve for

There may be formulas for introducing people to each other.

(12) 'I'd like to introduce you.'
Wà:kehre' *ki:*
I thought this
takwaterò:seron'…
I would make you all friends

Documenting such routine ways of interacting can be one of the most important contributions one can make to a community. Amery, working with the Kaurna, Yolngu, and Pintupi people of Australia, makes a similar point. He proposes what

he calls the Formulaic Method (2000: 209–212, 2001: 200–204, 2009) for learning a language that is no longer spoken:

> Well-formed, high-frequency utterances are learned, starting with minimalist utterances, such as stand-alone question words and one-word responses to questions that can be dropped into English conversation, but still maintain the grammatical integrity of the language. (Amery 2009: 139)

This approach can raise issues for communities to consider. Styles of interaction can change over time. Should traditional patterns of behavior be preserved, or should the heritage language be adapted to the modern world? There are communities, for example, in which people did not traditionally greet each other every time they met. Does this mean that new generations of learners, who have perhaps grown up greeting most people they encounter, should refrain from greeting others, or should new expressions be constructed in the heritage language to serve these purposes? In some societies, people have not traditionally thanked each other for small courtesies. Should learners refrain from thanking others in the heritage language in situations in which they would do so now in the majority language? Issues of this kind are also discussed in Hinton & Ahlers 1999 and Mithun 2007. Whatever decisions are made by communities, the documentation can be a useful resource for informed choices.

3. Convergence

The second type of language endangerment identified by Noonan comes through convergence with an encroaching language. For an adult second-language speaker, learning a heritage language which is structurally similar to the encroaching language is certainly easier. A focus on the similarities can make the task seem simpler. But if all languages are viewed as essentially equivalent, the value of each becomes marginal. Recognizing what is special about a language can be a powerful tool for building respect and esteem, within the community and beyond. Recognizing the specialness is crucial for linguistics as well. Linguistic theory teaches researchers to spot patterns that have previously been identified and discussed. But it is just as important to discover the unexpected. Languages can be special in a variety of ways, some immediately obvious, some more profound.

3.1 Distributing information over words

A basic but often unrecognized way languages vary is in how speakers package information. What speakers of one language may package in a string of words, for example, speakers of another may package in just one. Compare the Mohawk word in (13) with its English translation.

(13) Mohawk polysynthesis
Ahsani'tskwahra'tsherakarhátho'.
'You might tip over a chair.'

Not surprisingly, the Mohawk word has multiple meaningful parts.

(14) Morphological analysis
Ahsani'tskwahra'tsherakarhátho'.
aa-hs-an-i'tskw-a-hra-'tsher-a-karhat-ho-'
might-you-own-bum-LINKER-set-thing-LINKER-turn-cause-PFV
'You might cause the thing one sets one's bum on to flip.'

Does this difference matter? Indeed it does. Many of the distinctions Mohawk speakers slip inside of a word can also be expressed in a separate word. Alongside of the prefix *aa-* 'might, could, should' there are full words like *enwá:ton*' 'it will be possible'. In addition to pronominal prefixes like *-hs-* 'you' there are separate pronouns like *í:se*' 'you'. In addition to causative suffixes like *-ho-* there are whole verbs that can specify causation. Speakers often have choices about whether to package a cluster of ideas together within a single word, or to spread them out across a string of separate words.

The alternatives are not equivalent. What may be appropriate in one context may not be appropriate in another. One way might highlight one aspect of the message, and another leave it in the background. One way may be idiomatic and another simply not the way things are said. Mohawk speakers cannot generally identify parts of words consciously (unless they are also linguists), but they often manipulate them with dazzling skill. And they often have some sense of what is inside of their words. As one speaker remarked, 'The language paints pictures'. Referring to a certain lady, for example, one speaker used the expression in (15).

(15) *Teiakotia'ta'neká:ron*
te-iako-at-ia't-a-'nekar-on
apart-she-self-body-LINKER-explode-STATIVE
'She has bodily exploded.' = 'She was a large woman.'

As is well known, language is a rich repository of culture. The Mohawk verb *wa'khehárhahse'*, for example, means 'I bought her a gift'. It begins with the factual prefix *wa'-* for past tense and the pronominal prefix *khe-* 'I/her'. It ends with the benefactive applicative suffix *-hahs* 'for', and the perfective suffix *-e'*, which indicates that the event is viewed as a complete whole. The surprise is the root: *-har-* 'hang'. The whole word means literally 'I hung it for her'. Speaker Kaia'tit'ahkhe' Jacobs explained the history of the term. Traditionally when a couple married, a rope was strung in the house for guests to hang blankets on as gifts. The verb 'hang for' was extended to refer to gift giving in general.

3.2 Distributing information over categories

Much of the existing documentation of minority languages consists of translations from a contact language. The words themselves may be from the traditional language, but their arrangement often reflects lexical categories of the model. This phenomenon can be seen in the Bible passage in (16), translated from English by excellent, first-language Mohawk speakers.

(16) Nouns and verbs in Jonah 2:10.
E'thó' nene Roiá:ner wahrónnien ne kéntson
then the Lord he made the fish
'Then the Lord ordered the fish
ne taontahóstike' ne Jonah ne thiió'ke.
the he would spit him out the Jonah the beach place
to spit Jonah up onto the beach.'

The words are all indeed Mohawk (with the borrowed name *Jonah*). But it is immediately clear that this passage did not originate in Mohawk. Mohawk speakers would not have presented things in this way. The constituent order is a word-by-word replication of the English Subject-Verb-Object, not technically ungrammatical, but unlikely in this context. More telling is the density of referring expressions. Mohawk has just three kinds of words: nouns, verbs, and particles. Spontaneous speech in English shows a Noun/Verb ratio of approximately 1/1 (Wallace Chafe p.c). Similar counts in Mohawk show a Noun/Verb ratio of 1/17. The sentence in (16) contains four referring expressions: 'Lord', 'fish', 'Jonah', and 'beach'. Compare (17), from spontaneous conversation.

(17) Spontaneous Mohawk conversation

	VERB	PARTICLE	PARTICLE
A.	*Tesewatenna'tsherénhawe'*	*ken*	*ní:se'?*
	you lunch have brought	Q	yourself
	'Did you bring your lunch?'		

	PARTICLE	PARTICLE
B.	*Hen.*	*Kenh.*
	'Yes.	Here.'

	PARTICLE	VERB
A.	*Háo'ki*	*tetewatskà:hon.*
	come on	you all and I dine
	'Come on, let's eat.	

PARTICLE	PARTICLE	PARTICLE	VERB
Ó:nen	*ki:*	*ni'*	*wa'katonhkária'ke'.*
now	this	myself	I got hungry
Now I'm hungry.'			

VERB	PARTICLE	PARTICLE	VERB
Wakerhà:re'	*ki:*	*ni'*	*a:katshó:ri'.*
I am waiting	this	myself	I would slurp

'I'm waiting to have some soup.'

There is not a single noun in the entire interaction. The pattern is typical. A number of factors contribute to it. In some cases verbs contain nouns inside of them, like *-atenna'tsher-* 'lunch' in 'you have lunch brought'. In others, a verb alone captures the situation: the verb *-atshori* 'slurp' is used only for eating soup or something soupy. For eating something else, a different verb is used.

Another salient aspect of the exchange in (17) is the density of particles, especially in contrast with the sentence that originated in English in (16). Particles are pervasive in Mohawk, especially among highly skilled speakers. They convey all sorts of information, some grammatical, some pragmatic. Most particles are below the consciousness of speakers and difficult to translate. Interestingly, they tend to disappear when speakers write. (Even first-language Mohawk speakers learn to write first in English.) The result is that particles are typically sparse in language lessons, and, accordingly, not learned by second language speakers.

An example of the subtlety of particle use can be seen in the excerpt from a conversation below. Two people were discussing an accident in which a lacrosse player had been killed by a goalpost. One noted that the goalpost was constantly undergoing repairs. The next comment contained a barely perceptible particle *se'*.

(18) *Wa'ka'rhé:nien'ne' se' wáhi'.*
it toppled over TAG
'It fell, didn't it.'

The particle *se'* indicates that this statement contrasts slightly with the preceding. It is not as strong as English 'but' or 'however', which are rendered by other means in Mohawk. It can appear in conversation to indicate a slight difference in assumptions from those of the previous speaker. The comment in (18) also includes another particle *wáhi'*, a tag something like the English *Didn't it*. The Mohawk tag has a wide range of uses, some parallel to those of English tags, some not (Mithun 2012a). Like its English counterpart, it can solicit confirmation from the listener in cases of doubt. It can be used to draw a listener into the conversation. It can mark recognition of the knowledge of the listener. It can also be used to highlight an important point, as here, focusing on a particular idea and requesting acknowledgment from the listener. The listener's response is in (19).

(19) *Né:' ki' wáhi'.*
that in fact TAG
'It did indeed, didn't it,

ranontsì:ne.
on his head.'

The particle *ki'* is also very frequent in speech, though often barely audible. It signals that the remark is pertinent to something that has just been said. The response to this is in (20).

(20) *Thó* *takà:ra'ne'*.
there it came there to set on
'That's where it fell.'

The word *takà:ra'ne'* would have been a complete grammatical sentence on its own, with the direction of motion indicated by the prefix *ta-* 'to there'. The particle *thó* 'there' links this statement to the preceding, referring specifically to the location mentioned by the previous speaker, the boy's head.

Particles like these are pervasive in the speech of skilled first-language speakers, particularly in conversation. Their contribution can be subtle, but they play powerful roles in shaping the flow of ideas and conversation, creating coherence, linking contributions from the various participants, and facilitating interaction. They rarely if ever appear in translations from English, or sentences constructed in isolation, even by good speakers. They are often conspicuously absent from language curricula and the speech of second-language speakers.

Documentation of everyday conversation in natural contexts is crucial if the special properties of the language are to be appreciated, both by descendants interested in their own heritage, and by linguists seeking to understand the workings of a wide range of languages.

3.3 Idiomaticity

The point that there are many possible ways to say things, and that knowing a language well involves knowing which of them to use, has been made eloquently by Andrew Pawley (Pawley & Syder 1983; Pawley 1986, and elsewhere). Pawley notes, for example, that English speakers could in principle use any of the alternatives in the sets of sentences below, but they know that the first is the appropriate one.

(21) Pawley alternatives

a. *I'm so glad you could bring Harry.*
b. *That Harry could be brought by you gladdens me so.*
c. *Your bringing of Harry causes me to be so glad.*
d. *I am in a high state of gladness because you could bring Harry.*

(22) Pawley alternatives

a. *The time is twenty to twelve.*
b. *The time is eleven o'clock and two thirds.*
c. *The time is a third to twelve.*
d. *The time exceeds eleven by fifty minutes less ten.*

Documentation of spontaneous speech is crucial if such knowledge is to be preserved.

Mohawk morphology is tightly structured. It is templatic: the morphemes which can appear in each position within the word are rigidly defined. It might be thought that since the possible combinations of prefixes, roots, and suffixes are so strictly specified, simply following the rules would be sufficient. But first language speakers bring much more knowledge to speaking than those rules. One speaker made the remark in (23).

(23) *Iah énskak tsi, teió:ien' ne,*
not one only as not does it have the
'There's more than one way
a:kì:ron' ne iaesate'nikonhrón:ti'.
would I say the you would throw your mind
to express yourself.'

Speakers know which of the alternatives are appropriate for the context and their goals. They also know which combinations are idiomatic, traditional ways of expressing ideas. For the concept 'expressing oneself' he used the term 'throw one's mind'. The language is full of idioms built on the noun root *-'nikonhr-* 'mind'; they are generally terms that people know as part of the language, rather than create on the spot.

The loss of idiomaticity is characteristic of what Noonan described as convergence. When a group was discussing whether a friend had had her baby yet, a second-language speaker made the comment in (24).

(24) *Thé:nen' wakathónte'.*
nothing I have heard
'I haven't heard anything'

A first-language speaker later noted that a more usual way to express this fact would be (25).

(25) *Iáh tewakerihwarón:ken.*
not has it matter bumped into me

This does not mean that this second-language speaker's achievements in expressing herself are any less than admirable. Her fluency is impressive. But extensive documentation of language in use can provide models for the next level of expertise, idiomaticity.

The traditional way of describing someone with white or grey hair is in (26a). Now one often hears (26b).

(26) Grey hair

a. Traditional
Iakohrà:then.
iako-hra'th-en
F.SG.PATIENT-be.white.haired-STATIVE
'She has white or grey hair.'

b. Innovative

Karà:ken	*niiakononhkwiserò:ten.*
ka-rak-en	ni-iako-nonhkwiser-o't-en
N-be.white-STATIVE	PARTITIVE-F.SG.PATIENT-head-be.a.kind.of-STATIVE
white	so is she headed

'She has white hair.'

(26b) is a more literal translation of the English, the same general construction one would use to say 'the dress is white'. One speaker noted that even first-language Mohawk speakers now sometimes use this expression: "People are modeling their Mohawk on English, to try to make it easier for learners." Similar new usages can be heard for descriptions of people with blond hair or blue eyes.

Another innovative use can be heard in talk about speaking a language.

(27) Speaking a language

a. Traditional
Sahrónkha' ken?
'Do you speak?'

b. Innovative
Satá:tis ken?
'Do you talk?'

Both verb roots, *-ahronk* and *-atati*, are native. The first, *-ahronk*, is the usual way to describe the ability to understand and speak a language. The second, *-atati*, simply means 'talk'. The use of *-atati* to ask whether someone can speak Mohawk is a calque on the English *speak*. A first-language speaker remarked, "They should know better. They're trying to help the non-fluent speaker understand what they're saying, then it becomes a habit."

Innovations like these pose choices for communities. Convergence, like that seen in the two examples above, is natural with bilingualism. As noted earlier, bilingualism is empowering, allowing speakers to participate in two cultures and to talk with a wider range of people. Bilinguals have more options at their disposal for expressing their ideas. At the same time, bilingualism can affect both languages in the ways speakers package ideas. It is up to communities to decide whether or not to try to influence such effects. Documentation of spontaneous speech can

help them become aware of the kinds of processes that occur and make informed decisions as they work toward revitalization. And it can shed light on the ways contact can shape the development of languages in general.

4. Structure and substance

Linguists are naturally attracted to recurring structures and patterns. General principles help bring order out of apparent chaos, and provide learners with powerful tools for mastering the language. But the strong connections among structure, substance, and use are becoming ever clearer. Languages are rich, dynamic systems, constantly evolving as speakers routinize recurring expressions and extend existing patterns creatively. Good documentation of spontaneous speech, where speakers are the ones to choose both the grammatical structures and the vocabulary attached to them, and where the contexts can be seen, can be important for language learners hoping to speak idiomatically. It is also important for our understanding of the processes that shape language.

4.1 Routinization

An example of routinization can be seen in the development of the Mohawk negative construction seen earlier. Basic negative constructions are formed with the particle *iáh* 'no, not' plus a negative prefix such as *te'-* on the following verb: *Iokennó:ron* 'It is raining', *Iáh te-iokennó:ron* 'It is not raining'. To negate just a part of a statement rather than the whole fact, a special construction is used based on the verb root *-i* 'be'.

(28) Negation
Iáh è:rhar tè:ken
iah ehrhar te'-ka-i
not dog NEGATIVE-NEUTER-be
not dog it is not
'It is not a dog.'

This construction can be seen in (29) 'It is not Indian', where the term *onkwehonwehnéha'*, literally 'real person style', used here for referring to the Mohawk language, is a nominal.

(29) Negation in speech
Iáh tetkaié:ri,
not is it right
'It's not correct.

iáh onkwehonwehnéha' tè:ken.
not real person style not is it
It's not Indian.

Iáh nonwén:ton tewakathontè:'on
not ever have I heard
I'd never heard

tsi iáh onkwehonwehnéha' tè:ken …
that not real person style not is it
that it wasn't Indian...'

Very often, however, the final syllable of this frequent construction is omitted: *tè:ken* > *tè:*.

(30) Shortened negation
Iáh né: tè:.
iah nè:'e tè:ken
not it is not is it
'That wasn't it.'

Here the verbal structure is gone; both the pronominal prefix *ka-* and the root *-i* 'be' (*ka-i* > *ken*) have simply been dropped. What we now have is an unanalyzable particle *tè:*. Since both the full form *tè:ken* and the short form *tè:* still occur, we can see the development of the language in action.

Particles themselves may become further reduced with frequent use. The particle *shes* 'customarily, habitually, formerly' is often shortened simply to *s* in certain combinations in rapid speech. Both forms can be seen in (31).

(31) Shortened particle
Thos wáhe' thoió'te'.
thó shes wáhe' thoió'te'
there customarily TAG there he worked
'He used to work there didn't he.

Tanon' shes ki: ratiksa'okòn:'a tsi nihonwáhsons.
and customarily this children so they hate him
And the children really used to hate him.'

Of course not all verbs evolve into grammatical or discourse particles, and not all particles are reduced at a constant rate. Structure does not evolve independently of substance, and substance affects frequency of use. Documenting spontaneous speech in context can allow us glimpses of the relative frequencies of forms and the evolving uses to which they are put, helping us to understand how language structure develops over time.

4.2 Extending patterns

Grammatical constructions may be lexically-specific to varying degrees, that is, they may be more or less general, more or less tied to particular vocabulary. In his masterful survey of complementation, Michael Noonan (2007) noted that cross-linguistically, complementizers often cooccur with specific sets of matrix verbs. In English, for example, we have *I know that she'll read the book,* but *Joe wants Pete to retire*, and *Bert made Jimmy __ blush*. Mohawk shows similar differences.

(32) Mohawk zero complementizer
Wà:kehre' ki' __ takwaterò:seron' …
I thought in fact I would make you all friends
'I thought __ I would introduce you.'

(33) Mohawk *ne* complementizer
Tóka'ni' seweién:ton ne ohsnónhsa' ahsewennahnó:ten'.
maybe you know how COMP hand you could word stand
'Maybe you know how to read sign language.'

(34) Mohawk *tsi* complementizer
Waháttoke' ki' tsi rotihnekatárion.
he noticed in fact COMP they each had liquid inside
'He noticed that they had been drinking.'

Just one or two samples of Mohawk complement constructions would not provide sufficient information for a full understanding of the cooccurrence patterns. Learners need more than a few examples to equip them to create their own complex sentences with other verbs, and linguists need more in order to see the steps by which such constructions develop. Complement constructions often begin with a small set of lexical matrix verbs, and then are generalized gradually, item by item. There are clear patterns to the distribution of the three complementizer options in (32)–(34), but they are in a state of flux, varying across individual speakers and generations. The complementizer *tsi* appears to be slowly gaining ground in Mohawk, as speakers extend the contexts in which it is used. Rich documentation can often show us the routes by which such processes progress (Mithun 2012b).

4.3 Creativity: Language use

A significant part of the linguistic heritage of a community is what speakers do with their language. Mohawk speakers are known for their skill and delight in language, their propensity to play with it and use it creatively. One speaker mentioned that her grandmother used to refer to her best hat as *tsi kaná:taien' ieiakehtáhkhwa'*. It is immediately obvious that such a term must have more meaning than simply 'hat'. In fact it does.

(35) *tsi kaná:taien' ieiakehtáhkhwa'*
to it town lies away one goes with it
'I go to Cornwall with it'

The phrase *tsi kaná:taien'* literally 'where the town lies', is the term for the large town closest to that community, called *Cornwall* in English. The grandmother originally referred to her best hat jokingly as in (34), but thereafter, others around her picked up the term and used it for any fancy hat, tongue in cheek. Such expressions might not be the first thing that comes to mind when one is thinking about writing a grammar, but they are a part of the legacy of the language. They rarely surface under elicitation, but they can tell us much about how speakers exploit the structures available to them to create new expressions.

Mohawk speakers have a long tradition of valuing and cultivating language play. Speaker A below, was saying good-bye to B, whose Mohawk name is *Tekaronhió:ken*, literally 'Split in the Sky'. As a name, its primary meaning is to designate this man, and its literal meaning does not immediately come to mind when it is used. But another man, speaker C, jumped in, playing with the name.

(36) Linguistic virtuosity

A. *Ó:nen Tekaronhió:ken*
'Good-bye, Frank.
Kátkek tentsitsatátken.
Someday we'll meet again.'

B. *Hánio'*
'OK.'

C. [*Hé: Tekaronhiakháhsion.*
'Hey, Dividing the Sky.'

B. *Kwah nekne akoniahská:nekse'*.]
'I'll make you wish for it (seeing me again).'

C. *Oh naiá:wen'ne*
how could it happen
'How could
ne tekaronhió:ken ahsatóhetste'
the it is sky split you could pass through
you pass through a split sky
tóka' iáh tha'tekaronhiakháhsion?
if not is it sky divided
if the sky isn't divided?'

Documentation of language in use provides a record of speaker creativity that is an integral element of the linguistic heritage of the community. It also allows us insight into the ways speakers can manipulate the structures available to them.

5. Pride in complexity

Most of the structural complexity of a language is normally below the level of consciousness of first language speakers, particularly if the language is not written. It is in fact what makes them able to communicate so efficiently. As a result, however, the language can be undervalued by the very speakers who manipulate it so skillfully. One Mohawk group was commenting on the fact that they had been largely unaware of the richness and intricacy of their language until they began to write it.

(37) Complexity

A. *Iah ki' tetewattó:kas, nó:nen iáh teiokwahiatonhátie'.*
'We're actually not conscious of it when we're not writing.
Kwáh ionkwahiatonhátie' thó: ne: ó:nen.
'Just when we're writing, that's when.'

B. *Tewattó:kahs ki' tsi nikanontsistí:io's, wáhi'.*
'We realize how smart we are, don't we.'

A. *Tóka' ni iáh tekanontsistí:io's.*
'Or how <u>not</u> smart we are.'

Much of linguistic structure is easier to see in written form, where one can take time to examine patterns and contexts. Full documentation of active, spontaneous speech can substantially increase appreciation of and, accordingly, the potential longevity of the language.

At the same time, revitalization programs and grammarians alike struggle with reconciling full complexity and user-friendliness in the materials they produce. A group discussing the optimal level of complexity for a reference grammar for the community joked about a 'Mohawk for Dummies' version.

(38) *Wa'i:ron', "Mohawk for Dummies".*
'She said, "Mohawk for Dummies".

Hánio!
Come on!

Wa'i:ron' né: nè:'e,
She said, that one,

"*Iáh tè:kehre'", wa'i:ron',*
"I don't think so", she said.

That's an oxymoron!

The complexity, while sometimes daunting, can be an enormous source of pride. In the early 1970's, a group of dedicated and energetic Mohawk speakers began classes and workshops in Mohawk linguistics in preparation for teaching the language. As they and generations of teachers after them became conscious of the

enormous richness of the structure of their language, they constantly expressed wonder at the kinds of minds that shaped it.

(39) *Sewatié:ren's ò:ni' tiotenonhianíhton*
sometimes too it is frightful
'Sometimes it's frightful

tsi niió:re' tsi é:so' tekawennahsonterónnion'.
so it is far so many words are connected here and there
how many connections there are (morphemes within words).

Né: ki' aorì:wa' tho niió:re'
that is in fact its reason there so it is far
That's why,

tsi kanontsistií:io's ne Kanien'tkehá:ka.
so it is good heads individually the flint place people
Mohawks are so smart.'

Recently some of those in the first pioneering group commented that they felt that discovery of Mohawk morphology had been a major turning point in self-esteem for their community. A true appreciation of the systematicity and intricacy of the structure can come only with documentation of extended speech in context.

6. Conclusion

Fortunately, language documentation is now recognized as an important scholarly methodology in the field of linguistics. Among the points in the Resolution Recognizing the Scholarly Merit of Language Documentation passed by the Linguistic Society of America in 2010 is the following.

> **Whereas** 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;
>
> **Therefore** the Linguistic Society of America supports 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.

But the value of documentation is not limited to academia. What is recorded now may be all that is available to future generations. Rice (2011) makes the point that academic scholarship and community-based research need not be distinct endeavors, but that community-based research can yield traditional scholarly products as

well as new research topics. Yamada (2011), describing documentation and revitalization of the Kari'nja language in Suriname, shows how documentation and the creation of teaching materials support and strengthen each other.

Those engaged in revitalization projects can provide important advice on kinds of materials to include for posterity. Most express genuine gratitude for any and all documentation that exists of their heritage languages. But they are also discovering what materials have proven the most useful and what needs they have felt the most keenly.

A top priority is expressions for use in everyday interactions, words and phrases that learners can use early on and often. Next are model constructions that can allow learners to do more than name objects. Systematic grammatical patterns can help them to understand and create novel expressions from basic ones. Rich exemplification is useful: teachers need enough examples to create lessons, and students need enough to see generalizations for themselves.

But a language is more than a dictionary and abstract structure. Structure is intimately connected with substance and context, and these connections are part of the knowledge of first language speakers. Many perfectly grammatical structures are unidiomatic when combined with particular words and morphemes, simply not what is said. Chunks of language are also closely linked to context: speakers know what to say in particular situations. Much of the essence of a language has scope beyond the single sentence uttered by a single speaker. Particles that link ideas and give texture to interaction, for example, rarely occur in teaching materials or, for that matter, in the speech of learners.

All of these aspects of language are important both to scholars interested in the shapes that language can take and forces that mould them and to the people whose heritage the language represents. They are often most richly represented in documentation of extensive bodies of unscripted speech, in a variety of genres and contexts, with a strong interactive component. Documenting what speakers do when left to their own devices, in the varied contexts of their daily lives, can provide a foundation for appreciation of how each language is special and what makes it the way it is.

References

Amery, Rob. 2000. *Warrabarna Kaurna: Reclaiming an Australian Language*. Lisse: Swets and Zeitlinger.

Amery, Rob. 2001. Language planning and language revival. *Current Issues in Language Planning* 2(2–3): 141–221.

Amery, Rob. 2009. Phoenix or relic? Documentation of languages with revitalization in mind. *Language Documentation and Conservation* 3(2): 138–148. ⟨http://hdl.handle.net/10125/4436⟩

Hinton, Leanne & Ahlers, Jocelyn. 1999. The issue of "authenticity" in Californian language restoration. *Anthropology and Education Quarterly* 30(1): 56–67.

Linguistic Society of America. 2010. Resolution recognizing the scholarly merit of language documentation. ⟨http://www.linguisticsociety.org/about/what-we-do/resolutions-statements-guides/lsa-res-lang-doc⟩

Mithun, Marianne. 2007. What is a language? Documentation for diverse and evolving audiences. *Sprachtypologie und Universalienforschung* (*Language Typology and Universals*) 60(1): 42–55.

Mithun, Marianne. 2012a. Tags: Cross-linguistic diversity and commonality. *Journal of Pragmatics* 44: 2165–2182.

Mithun, Marianne. 2012b. How fascinating! Insubordinate exclamations. Presented at symposium on The Dynamics of Insubordination. Institute for Languages and Cultures of Asia and Africa, Tokyo.

Noonan, Michael. 2005. Language documentation and language endangerment in Nepal. Lecture, University of Iceland, Dialog of Cultures. ⟨http://archiv.ub.uni-heidelberg.de/savifadok/volltexte/2008/201/pdf/Iceland_Talk_Handout.pdf⟩

Noonan, Michael. 2007. Complementation. In *Language Typology and Syntactic Description*, Vol. 2, Timothy Shopen (ed.), 52–150. Cambridge: CUP.

Pawley, Andrew & Syder, Frances. 1983. Two puzzles for linguistic theory: Native-like selection and native-like fluency. In *Language and Communication*. J.C. Richards & R.W. Schmidt (eds), 191–227. London: Longman.

Pawley, Andrew. 1986. Lexicalization. *Georgetown University Round Table of Languages and Linguistics: The Interdependence of Theory, Data, and Application,* Deborah Tannen & James Alatis (eds), 98–120. Washington DC: Georgetown University Press.

Rice, Keren. 2011. Documentary linguistics and community relations. *Linguistic Documentation and Conservation* 5:187–207. ⟨http://hdl.handle.net/10125/4498⟩

Yamada, Racquel-María. 2011. Integrating documentation and formal teaching of Kari'nja: Documentary materials as pedagogical materials. *Language Documentation and Conservation* 5:1–30. ⟨http://hdl.handle.net/10125/4486⟩

Unanswered questions in language documentation and revitalization

New directions for research and action*

Lenore A. Grenoble
The University of Chicago

The last twenty years have witnessed an explosion of research on issues of language endangerment, with the emergence of documentary linguistics and the growth of language revitalization programs, resulting in changes in methodologies and in subfields within linguistics. The present article assesses this work in terms of its impact, focusing on documentation corpora, their contents, and how data are collected, archived, and used. The push to document the "last" fluent speakers has resulted in gaps in our current research, such as a general lack of documentation of variation, few studies of the kinds of change that take place during language shift and attrition, and few studies of the newer forms of language which emerge as the result of revitalization.

1. Taking stock

My comments here are an attempt to capture some general questions and even dissatisfactions that are in the air, emerging in both formal presentations and informal discussions after symposia like the one which have resulted in the present volume, or in a variety of publications which openly engage in these debates. See, for example, Evans (2008), Gippert et al. (2006), Himmelmann (2008), Woodbury (2011), as well as such journals as *Language Documentation & Conservation,*

* I would like to express my gratitude to the organizers of the Conference in Honor of Mickey Noonan at the University of Wisconsin, Milwaukee in October 2011, which stimulated lively discussion and thought on the topics presented in this paper. Noonan was well known and respected for championing the kind of linguistic research which would be directly relevant to language preservation and maintenance, and it is my hope that the thoughts presented here would be in keeping with his principles. Research was supported by funds from the Humanities Division at the University of Chicago. I am grateful to Jessica Kantarovich and Adam Singerman for their help with this research.

published at the University of Hawaii, Manoa; or *Language Documentation and Description*, published at the School of Oriental and African Studies, London.

In taking stock, I ask key questions about the nature and direction of current work in language documentation. Specifically, we need to consider not only what we are doing, but what we are not doing. What are we missing? What questions are we not asking?

Do we have the right tools? And do we have the right goals? These questions serve as the starting points for discussion. The answers in some cases may be no, but in others yes, that we are in fact doing the best job possible, given current conditions and constraints. Even if some of these questions turn out to be non-issues, it is still important to take stock of things, though we may decide we are doing the right things now. This is in part because the field has developed to a sufficient point that we need to ask these questions; in part because we have reached some twenty years since some of the early cries for work on endangered languages (Hale et al. 1992; Robins & Uhlenbeck 1991); and, in part because some of the key sources for funding language documentation are reaching their limits. (For example, the Dokumentation Bedroheter Sprachen (DoBeS) funding from VolkswagungStiftung has completed its final round, with the last round of applications due 15 September 2011.)

Before proceeding, however, it is important to consider the general situation in which we are presently working. Who is the "we" in language documentation? Language documenters include trained linguists and community members/speakers with little or no training in documentation. Interest in issues of language endangerment extends beyond this group to journalists, activists, and philanthropists. The Sorosoro Program ⟨http://www.sorosoro.org⟩, for example, involves participation from a wide group of people. Here I focus specifically on linguists and community members as those most directly engaged in documentation. Neither the linguistic community nor the community of speakers is homogeneous in its views and opinions and, furthermore, neither community is discrete. More specifically, it is important to bear in mind that not all linguists are alike. Documentation is interpreted differently by different linguists, and not all linguists who study endangered languages are engaged in documentation. In fact, not all linguists consider documentation to be important, although in general most linguists today agree that it is important to document endangered languages. By the same token, *speaker communities are hardly homogeneous.* More specifically, speakers differ in their attitudes toward language and in their commitment to revitalization. Critically, they differ in their levels of knowledge of their language. Furthermore, *neither group is discrete*: there are speakers of endangered languages who are linguists, and there are linguists who have become deeply embedded in communities, through marriage or other deep, long relationships, so the distinction between the

two is not always clear. In other words, there are linguists in speaker communities and speakers in linguistic departments. By the same token, there are community activists whose primary objective is revitalization but who are also engaged in documentation, just as there are linguists who focus on documentation but who are actively engaged in revitalization, too. My general sense is that the more we work together, the more the boundaries between these categories become blurred. Still, it is useful to keep in mind that the goals of these two different groups, to the extent that they are different, can be very separate.

It is difficult, if not impossible, to discuss language endangerment and documentation without attention to the needs of language revitalization. Moreover, such considerations are critical to successful documentation. This is not only because community members tend to be more interested in, and more engaged in, revitalization than in documentation. A consideration of the kinds of documentation needed for successful revitalization can shape the documentation itself in new and challenging ways. In order to determine this, we need good information about revitalization, not only what the goals of such programs are, but also how to assess the programs.

In what follows these questions are considered from different aspects of endangered languages, specifically, from the standpoint of goals, content, methodology, format, and what can loosely be called "broader impacts." These areas are all intimately linked but it is useful to consider them separately, in part because the way one is framed can have a direct impact on the other. What is required is not only new research areas, but also deep and multi-faceted collaborations (both intellectual and in practice) to rethink the goals and methods of documentation and revitalization.

2. Goals in documentation

Himmelmann (2008: 346) defines the goal of language documentation as providing "a lasting, multifunctional record of the linguistic practices attested at a given time in a given speech community and the knowledge speakers have about such practices." He emphasizes the importance of documenting both linguistic practices and metalinguistic knowledge (Himmelmann 2005: 8). Similarly, Woodbury (2011: 159) argues that "LANGUAGE DOCUMENTATION is the creation, annotation, preservation, and dissemination of transparent records of a language. While simple in concept, it is complex and multifaceted in practice because its object, LANGUAGE, encompasses conscious and unconscious knowledge, ideation, and cognitive ability, as well as overt social behavior; because TRANSPARENT RECORDS of these things must draw on conceptual understandings and techniques

from linguistics, ethnography, psychology, computer science, recording arts, and more […]."For Woodbury, then, documentation involves not only language, but also social practices; the two are inseparable.

One of the key questions which emerges from these definitions is exactly how broad the documentation should be, and how to put limits on the enterprise. How does one go about documenting all linguistic practices? One answer is to strive for a diverse corpus, but that is not sufficiently informative. Exactly what areas should be included in documentation and which can (or should) be omitted? There has been relatively little focused discussion about this issue, with a few recent exceptions (Evans 2008; Holton 2011). This speaks to the larger issue of relatively little *corpus theorization*, as defined by Woodbury (2011). Woodbury makes the point that language documentation is still documentation, regardless of whether the records "add up in some way"; corpus theorization is defined as "the ideas according to which a corpus is said to cohere or 'add up.'" Lacking rich discussion in corpus theorization, we lack means for evaluating a documentation corpus. When is a documentation sufficient? Complete? Is a complete documentation even possible? From the standpoint of funding agencies, there is pressure to document undocumented or "underdocumented" languages, and such agencies view languages which have been allocated more money as a lower priority for documentation than languages which have received less overall funding. Lacking an adequate theory of documentation, we have no real way of assessing the quality and sufficiency of documentation. Rather, we often turn to quantity and technical standards: sound recorded in uncompressed wav format at a rate of at least 44.1KHz, 16 bit, for example, is seen as "good" documentation. But this completely skirts the question of content. Corpus theorization would aid in answering the question of when a researcher should cease working on one language and move on to the next, if the goal of the research is "adequate" documentation. (I would note that this suggests that the only goal of research is documentation, and many linguists would disagree that one should move from language to language. In fact, that goal is directly at odds with maintaining a long-term commitment to a community. But these issues need discussion.)

The ever-increasing interest in language documentation over the past twenty years has meant a tremendous surge in the amount of documentation which is being done. Still, we tend to prioritize certain kinds of work over others, and to privilege languages which have a set of characteristics. These are, namely, high levels of endangerment, so that recording the last speakers is an urgent priority; features which make them of typological interest, including rare or unusual phenomena from the standpoint of what we currently know about the world's languages; and language isolates for their potential to inform us of historical linguistics. By definition, we thus exclude variation, preferring to document a

language meeting any of these characteristics over a variety (dialect, sociolect, or contact form) of an already documented language. These priorities are defensible, but they have far-reaching implications for the kinds of research we are engaged in, as well as for the questions we have thus implicitly chosen to ignore.

2.1 Which languages do we document?

To what extent do current practices in documentation achieve these goals? One result of the current emphasis on endangered languages is that perhaps less attention is being given to the documentation of *not* highly endangered languages or not immediately obviously endangered languages. This is not to say that this work is not being done at all, nor is it to say that only endangered languages are the object of study. Certainly, world languages with large numbers of speakers are the focus of considerable research; English is arguably the most thoroughly studied language today. There are, however, a great many languages which are not currently endangered, although they may be showing initial signs of shift. These are languages with smaller numbers of speakers than languages of wider communication, those in the middle range. Many are under pressure from majority languages and they might shift, or they might not. Possible examples from the regions where I work include Buriat (Mongolic; with 218,557 speakers in Russia, 2010 [ISO 639–3 bxr]; an estimated total of approximately 283,000 speakers for all varieties [ISO 639–3 bua]) and Sakha/Yakut (Turkic; with 450,140 speakers in Russia, 2010 [ISO 639–3 sah]), both under pressure from Russian. There are other reasons to argue against the urgency of documenting these languages, such as the fact that they belong to relatively large and relatively well-studied language families, for example, or the fact that we already have basic descriptions of them. At the same time we have limited studies of conversation, discourse and information structure. Moreover, a brief look at census data suggests that the number of speakers is growing for Sakha, although the figure dips from 2002 to 2010 (from 363,000 in 1993 to 456,288 in 2002 and to 450,140 in 2010), and is declining rather dramatically for Buriat (from 369,000 in 2002 to 218,557 in 2010).[1] Both are spoken in official Republics of the Russian Federation, regions with the largest degree of autonomy possible within the Russian federal system. Thus a focused study of their language ecologies and the measures the two different regions have undertaken to support

1. All-Russian Census data; 2012 statistics for language and speakers are available at ⟨www.gks.ru⟩; 2002 at ⟨www.perepis2002.ru⟩. It is difficult to know how to interpret the rise in Sakha speakers from 1993 to 2002 and equally difficult to interpret the decline in 2010. Russian census data rests on self-reporting. There is widespread agreement in Buriatia that the Buriat language is shifting.

their usage would be informative to others. If we base our decisions on which languages to prioritize for documentation solely on scientific criteria, specifically, in terms of typological interest, rarity of certain linguistic structures and so on, we might easily decide to not consider Turkic and Mongolic varieties. Yet studying these specific language ecologies might provide more insights into understanding both the social and linguistic process of language shift. The way we have formulated language documentation means that it provides little data about the very situations which have put these languages at risk.

This raises the next point: that another underdocumented area is the study of *languages in contact*. Of course, most endangered language situations are contact situations: most of today's languages are disappearing due to shift to another language or languages, not because the speakers are themselves disappearing. This is not to say that we do not study contact linguistics: there is a vast body of work on language contact, theoretical and descriptive, and numerous studies of the effects on linguistic structures due to contact. Yet there is very little documentation of contact in process, with the exception here being research on code-switching. But that work is largely grounded in sociolinguistics; few studies of code-switching stem from language documentation, and few documentation projects focus on code-switching. These gaps are due to several factors. One is the ongoing division between sociolinguistics (which studies code-switching, language variation, bilingualism and diglossia, and social networking – all topics of direct relevance to language endangerment) and language documentation. Second, an emphasis on documenting the last speakers of a language – even as problematic as the concept of the "last" speaker may be (Evans 2001; Grinevald & Bert 2011) – has serious repercussions in terms of the kind of documentation we do. It generally means that we may begin documenting only after the processes of shift and attrition are far along, and there is no longer a very large body of speakers who are bilingual, or nearly so, and in some earlier stage of shift. This means that we cannot document (and study) the beginning or middle stages of shift and language attrition. Moreover, there is a strong push to document fluent speakers, since the purported goal of documentation is to create a record of how fluent speakers use it. The net result, however, is an unintentional but clear kind of linguistic purism: non-fluent speakers do not "count" in the documentation, or are not the target of documentation, in an effort to document the last speakers. Not only does this potentially valorize both the variety spoken by one set of speakers as well as the speakers themselves, but it can actually marginalize other speakers and other varieties. In the extreme, these attitudes can further language shift, if speakers come to perceive the variety they speak as not being sufficiently pure or authentic to warrant documentation.

One consequence of this linguistic purism is that there are very few studies or documentation of *emergent languages*, languages which emerge as the result

of revitalization. These are in contrast to languages with uninterrupted transmission, learned from birth and carried on from generation to generation. Such languages are inevitably subject to change, and historical linguistics studies this change. But what of languages which are taught after interrupted transmission? These languages have much in common with ones that are taught as second or foreign languages but differ in a crucial respect: learners of the non-endangered language have access to a community of fluent speakers, while the learners of the endangered language often do not. In some communities the number of speakers who are willing and able to teach the language is very limited, and so revitalization in such cases often critically depends on first preparing a cadre of second-language learners to teach the language. It is thus often the case that adult second-language learners take up the task of teaching the language in the schools or elsewhere, as the people who learned the target language as children (and can thus claim first-language proficiency or fluency) are too elderly, too remote, and/or too few. For example, Tlingit revitalization in Alaska depends heavily on the elders who speak it as a first language to work with the teachers who have learned it as a second language. In the extreme, reclamation programs recreate the language on the basis of extant documentation, generally only written materials, often relying on comparative reconstruction using data from closely related and better described (or still living) languages. Prime examples are the reclamation programs for Myaamia (or Miami, ⟨www.myaamiafoundation.com/⟩) and for Wôpanâak (or Wampanoag, ⟨www.wampanoagtribe.net⟩); both languages were resurrected after generations of no speakers. The Myaamia Project for Language Revitalization and Cultural Awareness has relied heavily on the earlier work of linguists who documented the language while it was still spoken, and on the work of modern linguists to extract a description from that documentation (e.g. Costa 1994). In all of these cases, it is hard to imagine that the language spoken by the new generations of speakers matches the one that would have resulted from uninterrupted transmission from generation to generation, but how it differs is an open question. There are interesting scientific questions about the linguistic structures of such emergent languages and the impact of the first language of these speakers, but these questions remain largely unasked.

In sum, our very goal of documenting languages while they are still spoken has arguably led to a monolithic and limited view of what the goals and content of documentation projects should be. The push to record the best speakers, defined in terms of speaking the ancestral language most fluently, and the push to document languages with very few speakers, can both stem from and result in linguistic purism. The fact that we have very limited time and resources, coupled with the increasing requirements of documentary linguistics (see especially Evans 2008: 342–343), means that certain languages will not be documented, and certain

aspects of language use will not make it into documentary corpora. Under this view, no documentation can ever be complete. Regardless, purism in language documentation has resulted in a lack of research on how languages change in contact, on the very processes of attrition, and on fundamental questions as to differences in change due to shift and attrition on the one hand versus change due to stable contact on the other, or even whether such differences exist.

3. Goals in revitalization

It would seem to be a truism that the basic goal of language revitalization efforts is the creation of new speakers. Although this would most obviously appear to be the case, much centers around the issue of how "new speakers" are defined. On one end of the scale, the goal might be to create fully fluent, monolingual speakers who use the target language in all domains. For most, if not all, programs, this is an unrealistic goal. And perhaps on the other end of the scale is a primarily symbolic use of key phrases, sayings, greetings and possibly songs. Such symbolic use keeps the language present in the public sphere but does not entail real communication. In such cases it rather serves as an index for identity, but does not mean that speakers achieve, or even strive to achieve, fluency. Just which goals are appropriate depends on a nexus of factors, and different goals may be appropriate for different members of one and the same language community.

Thus setting appropriate goals in language revitalization is a complicated and ongoing task. Lindsay Whaley and I have argued that revitalization goals need to be realistic, determined with close attention to existing resources and to the overall willingness and commitment of those who will be engaged in revitalization efforts (2006: 21–49). This position is championed by Dauenhauer and Dauenhauer (1998: 62–63) who, following Fishman (1991), argue strongly for prior ideological clarification before embarking on language revitalization programs. Similarly, Grenoble and Whaley (2006: 160–176) call for assessment before beginning a revitalization program, on the basis that any group needs to know what the current situation is before determining how to change it. Such an assessment requires an evaluation of *resources: financial, language, and human*. *Financial resources* encompass the sources of money available within the community, the likelihood of obtaining external funding (from the federal government or humanitarian organizations for example), as well as the kinds of resources available for education and programming, the use of media, etc. *Language resources* include access to existing language materials, such as grammatical descriptions and dictionaries, textbooks, pedagogical materials, written and oral literatures, and so on. Moreover, language resources include available speakers of the target language who

can serve as teachers, create pedagogical materials, and so on. *Human resources* comprise the number of people who could be involved in creating and promoting language revitalization, and the skills they could bring to the process. This aspect of assessment necessarily includes language vitality and variation. It also includes the general level of interest of community members, both speakers and non-speakers, in using, teaching, and learning the language. Finally, it refers to the availability of experts, inside or outside the community, to assist in technical aspects of revitalization. In addition to all of these, it is important to determine the attitudes of the people potentially involved. First are the attitudes toward the local language versus one or more languages of wider communication. Second are the attitudes toward different variants of the local language. These are relevant in a host of ways in revitalization efforts: in choosing a basis for a standard or potential standard variety, it is important to assess these attitudes.

This discussion would seem to presuppose first complete assessment, and then the development of an overarching revitalization program before beginning any of the actual work. In reality this is often not the case. In many cases, multiple activities (assessment, resource development, teaching) begin at once. Still, many people report that the assessment process itself is instructional, and that often people are surprised to learn that their community is undergoing shift. Child acquisition of language is so natural that it is often taken for granted. Bona fide assessment of language vitality is critical in this regard.

Successful language revitalization depends on assessment in a variety of other ways. Revitalization programs themselves need to be assessed, to see if they are achieving their goals and using resources appropriately and if those goals need to be recalibrated. There has been little theorizing about revitalization as a specific endeavor and, concomitantly, little discussion of how to assess revitalization programs. There are at least two key parties which have a vested interest in such assessments: the funding agencies and communities themselves. Here the notion of community needs to be broken down into the various stakeholders: teachers and other educators, language learners as a whole, parents and teachers in school-based programs, community leaders, activists and linguists, to name just a few. Even very fundamental questions about assessment still need wide discussion, such as how we determine when a program is successful. What criteria are used? How do we know when to shift goals and refocus, and when to stay the course? When do we move from revitalization to maintenance?

This is not to say that there have been no attempts to assess revitalization; see Gordon (2009) for Anishinaabe adult immersion; Hornberger and King (1996) for Quechua programs; Johansen (2004) for a survey of a number of North American programs; King (1999) for Quichua; Nyika (2008) for minority languages in Zimbabwe; Person (2005) for Bisu in northern Thailand; Walsh (2005) for

Australian aboriginal languages; and Wetzel (2006) for a survey of needs assessments for Potawatomi. These references indicate the geographic and linguistic richness in the range of both revitalization and assessment. They are quite honest, speaking not only of strengths and successes but also of failures and challenges. Some rely on quantitative measures and others on participant-observation; some study school programs for children, others adult immersion programs. But much more could be done, as many programs are not assessed (or the assessment is not readily available) and many of the assessment studies center around individual case studies, so it is at times hard to determine whether the findings are generalizable to other situations or represent idiosyncratic, or local, specifics. More data are needed with both quantitative and qualitative information to provide a fuller picture.

Finally, there are larger meta-issues which can be addressed by evaluating existing revitalization programs. First, it would be useful to know if revitalization has had any effect on language pedagogy. With the exception of the Master-Apprentice model (Hinton et al. 2002), most models of language revitalization are school-based. (I include pre-school program and language nest models in this category.) School-based models run the risk of creating language usage in one and only one domain: the school. Programs generally focus on the creation of new speakers, overlooking the need to create new domains as an inherent part of creating language vitality.

One major issue is whether revitalization programs affect the language attitudes of people inside a speaker community. Anecdotally, they seem to have a positive effect on community members, improving attitudes and prestige, and perhaps no effect otherwise. Is this true? Can these changes in community attitudes (if they in fact exist) be leveraged in some way? Can external attitudes be changed? This is potentially very important: there is broad agreement that language attitudes and language prestige affect vitality and usage. Can community-internal changes in attitude and prestige be mobilized to have a similar effect on other populations? This is a critical question, as changing attitudes and prestige are fundamental to reversing the situations which led to language shift in the first place.

4. Uniting documentation and revitalization

If documentation were to be reformulated so as to be driven by the goals of language revitalization, how would that change the nature of documentation, if at all? In some ways, most certainly; in others, not so clearly. For example, a fundamental part of documentation is what is referred to as "core linguistic work" by Comrie (2007), who notes that documentation of a traditional language is

required even when a linguist is committed to work on language revitalization. He focuses his claims around the centrality of potential contributions to linguistic science, arguing specifically that "the revitalization of a language does not obviate the need for documentation of the traditional language, since a revitalized language may differ quite extensively from the traditional language to which it corresponds, in particular through the loss of precisely those distinguishing features that make the traditional language of such paramount importance to linguistics" (Comrie 2007: 34). This frames the need for basic description in terms of advancing linguistic science. It is possible to conceive of another view, one that is not necessarily at odds with the needs of linguistic science but rather frames the issues from the standpoint of revitalization. For revitalization, core linguistic work is essential if the language is undescribed, or even underdescribed: learners need to know phonology, morphology, syntax, semantics and more. (This presupposes that the learners are not children learning it from birth but acquiring it as a second language.) They need to know the spirit of the language, how these pieces come together in actual use to create real communication. In this view, descriptive and documentary adequacy is achieved when a learner can extract sufficient information from the documentation to become a fluent speaker. Note that this provides an answer to the problematic question of when a documentation is complete, and further note that most (if not all) documentations are incomplete by this measure.

Beyond the need for basic descriptive work which can form the basis of revitalization materials – dictionaries, descriptive and pedagogical grammars, textbooks and exercises – most speakers are interested in how language is used. This argues for documentation which focuses on communicative practices, conversational structures and discourse, and varying registers and genres. To the extent that speakers desire to speak the language of their ancestors, they are interested in linguistic practices, in language which is situationally and socially anchored. Here I focus on bi- or multi-lingual situations of language shift, where use of the ancestral language tends to be limited to a few domains, primarily in the home.

To that end, documentation of contextually-situated language use is beneficial to linguistic theory and language revitalization alike. This includes linguistic etiquette, daily routines, and politeness mechanisms. All of these are directly relevant to people wishing to learn how to speak a language, and knowledge of culturally appropriate ways of verbal interaction. Note that this kind of documentation is also relevant to the study of semantics and pragmatics, not just anthropological linguistics. Similarly, the documentation of conversation is critical. Broad documentation of differing linguistic practices requires not only linguists, but other kinds of specialists who can understand and contextually situate such language

use. Potential partners include, for example, musicologists for understanding, interpreting and even accurately transcribing song and dance. Ethnobotanists are essential to the study of indigenous plant knowledge: they can identify plants and relate Western knowledge to local knowledge about plant habitats, uses and histories. They collaborate with linguists to understand the indigenous taxonomies, some of which are accessible through the plant names themselves (Si 2011; Whitecloud & Grenoble 2014). In Arctic regions, for example, sea ice engineers partner with linguists, anthropologists, geographers and Inuit hunters and fishermen to study changes in sea ice and climate. Krupnik et al. (2010) present compelling examples of the need for such collaborations in order to understand the multiple facets of environmental and social changes. These are just two examples out of a multiplicity.

Yet another avenue of linguistic research of benefit to revitalization is the study of languages in the early stages of shift. Documentation of the full range of register use is still possible, and such languages could presumably provide evidence for variation among fluent speakers as well as among shifting speakers, raising completely unanswered (and often unasked) questions about the latter category. Such documentation could serve as a pre-emptive social measure against further shift, prompting communities to recognize its incipience and prodding them to action. From the standpoints of both linguistic science and revitalization, it makes no sense to wait until languages shift to start documenting them. Even if our focus is shift and attrition, we can understand shift better if we know about pre-shift, conditions prior to shift. Current funding and research priorities rest on an implicit value system: we value more highly endangered languages over understudied but robustly spoken languages. If we actually care about revitalization and maintenance, we would be well-advised to study languages with smaller numbers of speakers that are not (yet) shifting. This may sound like pure heresy but, as Himmelmann (2008: 343) argues, the push to document the most endangered linguistic varieties "is obviously counterproductive and demoralizing when seen from the point of view of speakers and communities struggling to maintain their heritage language." I am not advocating that we abandon all work on endangered languages; obviously, that is very important work. But research on understudied and as-of-yet-unendangered languages is also imperative.

Finally I note that this claim – that we should expand documentation to other understudied, pre-shift speech communities – is based on the belief that the description and documentation of languages and linguistic structures is itself an incomplete and insufficient enterprise; our real goal should be a description and documentation of language ecologies. That is, we need to study languages as they are culturally and socially situated, in a full social context of production and use. And frankly, they always are. Abstractions out of that context are the products of

linguists, not speakers. I am not sure that we often achieve the goal of studying language ecologies and not just languages. I am not even sure what it looks like. Recordings and documentary corpora are essentially flat and linear. Actual language usage is multi-dimensional and multi-modal. How do we capture that? How do we transcribe and analyze it?

5. Conclusion

In the present article I focus on overall trends in language documentation and revitalization. In so doing, I ignore exceptions to these generalizations. That said, there is a general sense that it is time for documentary linguistics to further develop a theoretical basis to advance it beyond the definitions outlined by Himmelmann (1998, 2008), and a general sense that we need fresh conversations about how linguists and communities might better collaborate. One is for the intersection of community and research efforts; revitalization is considerably more challenging with more highly endangered languages and therefore not the place to begin.

In sum, I advocate that we work hard to avoid (intentional and unintentional) linguistic purism. We need to rethink some of our research targets and document not just ancestral languages but languages in the process of shift; to work with all kinds and varieties of speakers, at all levels of proficiency, and in all kinds of situations. We need to engage in lively, ongoing discussions and debate about corpus theorization and theories of documentation. Finally, we linguists need to aim for new collaborations, which will drive documentation from different viewpoints and will incorporate different research methodologies.

References

Comrie, Bernard. 2007. Documenting and/or preserving endangered languages. In *The Vanishing Languages of the Pacific Rim*, Osahito Miyaoka, Osamu Sakiyama & Michael E. Krauss (eds), 25–34. Oxford: OUP.

Costa, David J. 1994. The Miami-Illinois Language. Ph.D. dissertation, University of California, Berkeley.

Dauenhauer, Nora Marks & Dauenhauer, Richard. 1998. Technical, emotional, and ideological issues in reversing language shift: examples from Southeast Alaska. In *Endangered Languages: Current Issues and Future Prospects*, Lenore A. Grenoble & Lindsay J. Whaley (eds), 57–98. Cambridge: CUP.

Evans, Nicholas. 2008. Review of Gippert et al. 2006. *Language Documentation & Conservation* 2: 340–350.

Evans, Nicholas. 2001. The last speaker is dead – long live the last speaker! In *Linguistic Fieldwork*, Paul Newman & Martha Ratliff (eds), 250–281. Cambridge: CUP.

Fishman, Joshua A. 1991. *Reversing Language Shift*. Clevedon: Multilingual Matters.

Gippert, Jost, Himmelmann, Nikolaus P. & Mosel, Ulrike (eds). 2006. *Essentials of Language Documentation*. Berlin: Mouton de Gruyter.

Gordon, Christopher J. 2009. Anishinaabemowin Pane: A Qualitative Exploratory Case Study of an Indigenous Language Immersion Revitalization Program. Ph.D. dissertation, Cappella University.

Grenoble, Lenore A. & Whaley, Lindsay J. 2006. *Saving Languages*. Cambridge: CUP.

Grinevald, Colette & Bert, Michel. 2011. Speakers and communities. In *The Cambridge Handbook of Endangered Languages*, Peter K. Austin & Julia Sallabank (eds), 45–65. Cambridge: CUP.

Hale, Ken, Krauss, Michael, Watahomigie, Lucille J., Yamamoto, Akira Y., Craig, Colette, LaVerne, Jeanne Masayesva & England, Nora C. 1992. Endangered languages. *Language* 68: 1–42.

Himmelmann, Nikolaus P. 2008. Reproduction and preservation of linguistic knowledge: Linguistics' response to language endangerment. *Annual Review of Anthropology* 37: 337–350.

Himmelmann, Nikolaus P. 2005. Language documentation: What is it and what is it good for? In *Essentials of Language Documentation*, Jost Gippert, Nikolaus P. Himmelmann & Ulrike Mosel (eds), 1–30. Berlin: Mouton de Gruyter.

Himmelmann, Nikolaus P. 1998. Documentary and descriptive linguistics. *Linguistics* 36: 161–195.

Hinton, Leanne, Vera, Matt & Steele, Nancy. 2002. *How to Keep Your Language Alive: A Commonsense Approach to One-on-one Language Learning*. Berkeley CA: Heyday Books.

Holton, Gary. 2011. "Unknown unknowns" and the retrieval problem in language documentation and archiving. *Language Documentation & Conservation* 5: 157–168.

Hornberger, Nancy H. & King, Kendall A. 1996. Language revitalisation in the Andes: Can the schools reverse language shift? *Journal of Multilingual and Multicultural Development* 17(6): 427–441.

Johansen, Bruce E. 2004. Back from the (nearly) dead: Reviving indigenous languages across North America. *The American Indian Quarterly* 28(3–4): 566–582.

King, Kendall A. 1999. Language revitalisation processes and prospects: Quichua in the Ecuadorian Andes. *Language and Education* 13(1):17–37.

Krupnik, Igor, Aporta, Claudia, Gearheard, Shari, Laidler, Gita J. & Kielsen Holm, Lene (eds). 2010. *SIKU: Knowing Our Ice. Documenting Inuit Sea-ice Knowledge*. Dordrecht: Springer.

Nyika, Nicholus. 2008. 'Our languages are equally important': Struggles for the revitalisation of the minority languages in Zimbabwe. *Southern African Linguistics and Applied Language Studies* 26(4): 457–470.

Person, Kirk R. 2005. Language revitalization or dying gasp? Language preservation efforts among the Bisu of Northern Thailand. *International Journal of the Sociology of Language* 173: 117–141.

Robins, Robin H. & Uhlenbeck, Eugene M. (eds). 1991. *Endangered Languages*. Oxford: Berg.

Si, Aung. 2011. Biology in language documentation. *Language Documentation & Conservation* 5: 169–186.

Walsh, Michael. 2005. Indigenous languages of Southeast Australia, revitalization and the role of education. *Australian Review of Applied Linguistics* 28(2): 1–14.

Wetzel, Christopher. 2006. Neshnabemwen Renaissance: Local and National Potawatomi Language Revitalization Efforts. *The American Indian Quarterly* 30(1–2): 61–86.

Whitecloud, Simone S. & Grenoble, Lenore A. 2014. An interdisciplinary approach to documenting plant knowledge in Greenland. *Arctic* 67(1).

Woodbury, Anthony C. 2011. Language documentation. In *Cambridge Handbook of Endangered Languages*, Peter K. Austin & Sallabank, Julia (eds), 159–186. Cambridge: CUP.

Training as empowering social action

An ethical response to language endangerment

Carol Genetti & Rebekka Siemens
University of California, Santa Barbara

A recent response to language endangerment has been the rise of training programs in language documentation and conservation. Here we consider the position of training activities within the sociology of language documentation and conservation (LDC) work, specifically focusing on paradigms of ethical research and the relationship between academic and community partners. We examine one training program in depth, InField 2008, which had two distinct components: a set of short workshops, and intensive field training that incorporated language research. Grounding the discussion in the social science literature, we argue that training constitutes empowering social action and that different types of training promote different dimensions of empowerment. Training programs can be organized into a typology that is independent from, but overlaps with, the typology of research paradigms.

1. Introduction

Language contraction and loss is a process that has affected communities for millennia. Global historical trends have accelerated these processes.[1] Harmon and Loh (2010) demonstrate that the world's linguistic diversity declined twenty percent overall between 1980 and 2005, and that this contraction was especially realized across Indigenous languages in North America and Australia. Although language loss had been accelerating for decades, the extent and rapidity of the decline, and the profound influence that it would have on the world's cultural and linguistic diversity, only came to be widely recognized by linguists in the late 1980's and 1990's.

1. This article was largely produced when Carol Genetti was in residence at the Cairns Institute, James Cook University, in Cairns, Australia. Special thanks are due to Sasha Aikhenvald and RMW Dixon for their stimulating discussion, and to Komla Tsey, Janya McCalam, and Cath Brown for broader perspecive and deep insights born of decades of empowerment work in the health professions.

One of the primary responses of the field of linguistics to the prospect of widespread language loss was to prioritize *language documentation*, the creation of lasting multi-purpose records of the world's languages (Himmelmann 1998, 2006; Woodbury 2003, 2011). The past twenty years have seen a tremendous explosion of work in this field, with countless conferences, grant programs, archives, and manuals providing critical resources to field linguists. The theory and practice of language documentation have both developed significantly, especially with the advent of new technologies, helping this to be a remarkably productive effort that has filled archives with vast collections of high-quality digital language data that can serve as resources for heritage communities – and for academics – far into the future. Of course, there are still many languages without even simple descriptive materials, especially in some geographic areas (including the so-called language hotspots, "concentrated regions of the world having the highest level of linguistic diversity, the highest levels of endangerment, and the least-studied languages" (Living Tongues Foundation website, ⟨http://www.livingtongues.org/hotspots.html⟩)).

A second response to language endangerment has been increasing work in *language revitalization*, a term that refers to any activity "attempting to bring back endangered languages to some level of use within their communities" (Hinton 2011: 291). The types of activities that constitute language-revitalization programs are extremely diverse, including, for example: bilingual education, immersion schools, and other curricular programs for children; Master-Apprentice programs that pair elder speakers with young adult learners; Breath-of-Life programs that bring linguists and community members together to explore materials in archives; adult language classes, immersion summer camps, and the development of digital resources such as interactive websites and cell-phone apps. We use the term *language conservation* as a cover term that encompasses all activities aimed at strengthening the world's linguistic diversity, including language revitalization and the maintenance of minority languages that might otherwise decline.[2]

---

2. As in other fields in linguistics, terms regarding the variety of activities aimed at strengthening linguistic vitality are numerous and variously described and differentiated. Tsunoda (2006: 168–169) provides an overview of the terms used prior to the date of publication, with bibliographic references. Our use of 'language conservation' most closely matches his use of 'language revitalization', a term that we reserve for the more narrow focus of Hinton (2011), quoted above. The term 'conservation' arose through conversations at the University of Hawaii in 2005 and 2006 (Kenneth Rehg, personal communication). It was intentionally chosen as a cover term for the broad range of activities that have arisen in response to language endangerment, and is seen as being parallel to how the term 'conservation' is used in the environmental sciences.

A third response to language endangerment has been *training* in language documentation and conservation (Jukes 2011). This term is used broadly to cover a diverse array of activities, from one-on-one instruction to short workshops, small classes, intensive institutes, and formal degree programs. In its simplest conception the term 'training' refers to the transfer of skills and knowledge from one person to another. Although many people may primarily associate this term with skills development around technologies, in the context of language conservation, the scope of training is much broader. The transferred skills range from orthography development to filmmaking to linguistic analysis, and so on. The term can be thought of as referring more generally to *capacity development*, which provides communities with a range of skills and resources to most effectively realize their language goals. Training in language documentation and conservation can also be viewed as *education* in the broad sense of enhancing perspective and breadth of understanding, and frequently has significant personal impact on sense of self.

Language documentation and language revitalization can be seen as mutually reinforcing activities. The materials produced through language documentation can be used as input to language revitalization programs, while successful language revitalization creates contexts for language use that can be documented. This includes the documentation of the language revitalization itself, and the acquisition processes of those involved in language revitalization or reclamation. Training underlies both of these activities, providing community members and those that support them with the resources and expertise to be maximally effective. Training can thus be seen as a third axis in the effort to preserve linguistic diversity, both drawing from and reinforcing work in language documentation and revitalization.

The goal of this article is to consider the role of training activities within the sociology of language documentation and conservation (LDC) work. Many LDC projects involve both community-internal members and community-external academics, and there has been a remarkable increase in the ethics literature on the relationship between these actors in the conduct of LDC projects. The primary focus of this literature is on research that is physically situated within communities (i.e. fieldwork), and there are increasingly calls for *participatory community-based research* (e.g. Wilkins 1992; Furbee & Stanley 2002; Grinevald 2003; Yamada 2007; Penfield et al. 2008; Beier 2009; Benedicto et al. 2009; Cranmer et al. 2009; Michael 2009; Rice 2009, 2010, 2011a, 2011b; Guérin & Lacrampe 2010; Leonard & Hanes 2010). The question for the current discussion is how training activities fit within these models of community-academic collaborations and how our typology of such interactions can be expanded to include programs of this type.

This article will examine one training initiative in detail, the first Institute on Field Linguistics and Language Documentation (InField) that was held at

the University of California, Santa Barbara in 2008. Both authors were closely involved in the design and implementation of this institute, as Director (Genetti) and Assistant Director (Siemens). One of a number of new international training programs in LDC, InField was unique in bringing together people from both academic and community groups for shared training, thus bridging the academic-community divide. It is thus a lens through which to examine community-academic relationships. One of the explicit aims of InField was to foster understanding among participants and to advance the participatory community-based model of LDC, and there were a number of decisions made in the design of the institute to realize these aims.

Although InField promoted participatory community-based research, it was not itself a research project. Training differs from research in that its primary goal is transfer of skills and knowledge, and not the production of new knowledge through inquiry. It is important to note that research may be used as the medium through which training occurs, so training programs can be ranged along a continuum regarding the degree to which research is involved; however, even in 'research-heavy' training programs, the focus is on the transfer of skills and knowledge and the research itself is not the primary aim. In providing members of endangered-language communities with skills to advance the process of language conservation for the benefit of their communities, training in LDC constitutes *social action.* In addition, InField shared with the research models it promoted the feature of being *empowering*, a fact attested to by both participant comments and by community activities following the institute. InField, and similar training programs in LDC, can thus be conceptualized as *empowering social action*. Since language conservation is inherently a community activity realized through personal choices made by members of the speech community in the intimate and broader social contexts of their daily lives, it is only through the empowerment of speech communities that languages will continue to be spoken. Training programs, as social actions fostering empowerment, are thus likely to be critical to community efforts and to the preservation of linguistic diversity.

Section 2 provides a brief overview of discussions on ethics, empowerment, and research in the social sciences and in the endangered-language literature. A discussion of training in relation to these themes of the literature is presented in Section 3, viewed through the lens of InField. This section especially examines instances where InField was empowering for particular individuals, and will demonstrate that this occurred most profoundly in work that was strongly research-based. In concluding the article, Section 4 returns to the typology of ethical research paradigms and the relationships between research, training, empowerment and social action.

2. Ethics, empowerment, fieldwork models, and revitalization

One of the notable developments of the field of language documentation over the last twenty years has been a surge of interest in ethics, particularly as it pertains to the fieldwork context. This can be seen not only by the significant increase in scholarly articles on this topic, but also by the inclusion of explicit discussions of ethics in recent manuals on linguistic fieldwork (e.g. Dimmendaal 2001; Dwyer 2006; Mosel 2006; Crowley 2007; Bowern 2008; Chelliah & DeReuse 2010). The literature covers a range of topics, such as payment of consultants, issues of informed consent, access to data, and intellectual property rights. The majority of the work, however, focuses on the nature of the relationship between the community and the linguist, the complex power dynamics that underlie it, and the movement toward collaborative models of engagement which, in Dwyer's words, "entail indigenous people and field researcher's mediating each other's cultural imperatives…through productive mutual negotiation at the local level" (2006: 32).

The model of ethical fieldwork on endangered languages that has evolved over the last twenty years focuses on the linguist working for the community to meet the community's goals and on the community having control over the language, the collected materials, and programs for language maintenance or revitalization (Rice 2006; Czaykowska-Higgins 2009). The trajectory of this work can be traced back to the early 1990's (e.g. Cameron et al. 1992; Wilkins 1992; Craig 1993) and reflects the development of models of community-based participatory action research found in other social sciences (e.g. sociology, education, and the health sciences). Minkler and Wallerstein (2008) trace the intellectual roots of these models to two very different historical traditions: collaborative utilization-focused research and "openly emancipatory research, which challenges the colonizing practices of positivist research and political domination by the elites" (2008: 28). In line with the latter tradition, models of participatory action research reflect a growing awareness and appreciation of the agency and authority of traditionally marginalized peoples in research related to their selves, their communities, and their cultures.

Underlying the discussions on ethics in endangered-language research is a tripartite distinction made in the seminal work by Cameron et al. (1992, 1993). They distinguish three types of ethical research; these differ in how the relationship between the researcher and the research 'subjects' is conceptualized:

(1) Three types of ethical research identified by Cameron et al. (1992)
- 'Ethical research' *on* subjects
- 'Advocacy research' *on and for* subjects
- 'Empowering research' *on, for and with* subjects.

Whereas the traditional style of linguistic fieldwork would most likely be classified as 'ethical research' in this model, many field linguists also become involved in 'advocacy research' (especially with regards to land claims or other issues involving negotiations of community rights). In recent years, there has been a clear move towards 'empowerment research' on endangered languages. A definition of empowerment research from the social-science literature is that of Wallerstein (1992):

> A social action process that promotes participation of people, organizations, and communities towards the goals of increased individual and community control, political efficacy, improved quality of community life, and social justice. (Wallerstein 1992: 198)

Social action is broadly conceptualized as action that puts into practice convictions based on deeply-held values, frequently directed towards bringing about one of several alternative ends (Weber 1978: 24–25). The idea has come to include the concept of collective action that transforms the nature of local and broad societal conditions, as can be seen, e.g. in Horvath's definition of social action as "participation in social issues to influence their outcome for the benefit of people and the community" (1999: 221). By these definitions, community-based research that strengthens endangered languages constitutes social action, as it involves collective action towards a chosen social end based on community values. However many other types of linguistic research (for example, historical-comparative work, acoustic phonetic studies, or syntactic theoretical analysis) might not constitute social action in this sense, if it does not involve collective action by stakeholders around a social issue. Linguists (and other academics) whose work has social action dimensions are frequently motivated by a desire to use their expertise to positively impact society.

Research that results in the empowerment of social subjects has a number of hallmark characteristics (e.g. Minkler & Wallerstein 2008: 9; Tsey 2010: 9):

- It is participatory and cooperative, with community members and researchers jointly engaged and equal contributors within mutually respectful strategic partnerships;
- Participants negotiate the goals, methods, and outcomes of the research;
- It is grounded in relevant local knowledge traditions;
- It is capacity-building, in that it enhances and enables community initiatives;
- It thus achieves a balance between research and social action.

Within the literature on endangered-language documentation, many researchers promote research models that are empowering in some or all of these senses.

Crucially, these models locate the primary decision-making in community partners, who profoundly shape the goals and outcomes of the project, as well as negotiate the ownership and dissemination of research results.

One model of this type is Community-Based Language Research (CBLR). Proposed by Czaykowska-Higgins (2009), CBLR extends the tripartite distinction of Cameron et al. to a fourth level. She states:

> Community-Based Language Research, as I define it here, not only allows for the production of knowledge on a language, but also assumes that that knowledge can and should be constructed *for, with, and by* community members, and that it is therefore not merely (or primarily) for or by linguists. In a model of this kind, linguists are not the sole researchers, nor are they necessarily the leading researchers; rather, they are partners working in a collaborative relationship with members of the language-using community. (Czaykowska-Higgins 2009: 17; emphasis added)

Collaboration is explicitly articulated as a key concept and underlies this and similar models of field linguistics. CBLR, and similar research models, respond to the ethics challenge by placing control over the language, the research goals, data, and research outcomes directly into the hands of the language community. Like participatory action research in other fields, this model overtly addresses the power imbalances between researchers and communities (Rice 2006).

It is worth noting that much of the research that advocates for a fully collaborative ethical model has been based on interactions with endangered-language communities in North America or Australia. It is not clear how to extend such models – or whether it is appropriate to do so – to linguist-community interactions in other parts of the world. As Holton (2009) illustrates, ethical orientations are locally determined, and what is considered ethical behavior by a researcher in one community might be considered shockingly unethical in another. Given the relativity of ethical orientations, it is worth noting the fluidity of the notion of 'empowerment' itself, and the influence of the historical, ideological, cultural, educational, and socioeconomic contexts on how the idea of 'empowerment' can be understood and realized in different communities. An especially instructive example of this is described by Dobrin (2008), who reports on her experiences in Papua New Guinea. She notes that while the Western view of empowerment involves autonomy and self-determination, thus making it appropriate to place control over a project in community hands, the Melanesian view of empowerment reflects an individual's or community's ability to elicit exchange; the level that one participates in exchange relationships is a reflection of one's power. These two views are inherently in conflict, one being based on independence and autonomy and the other on dependence, as realized through exchange relationships and

networking. Dobrin's study illustrates the importance of situating the notion of 'empowerment' within local culturally-based systems of meaning (see also Nevins 2004). Thus, to be empowering, linguistic fieldwork and the development of language programs need to be sensitive to the complex ethical and cultural context in which they are based.

Regardless of the broader cultural context, when language revitalization programs are successful, they are empowering in the sense of Wallerstein (1992) presented above. Language revitalization is necessarily participatory, as it must involve at least a subgroup of community members who actively work to increase use of the language. Language revitalization programs are frequently grounded in local knowledge traditions, incorporating local songs, rituals, greetings, traditional cultural activities, and child rearing. Language revitalization programs are also designed to develop capacity, in that they build community structures supportive of language use. Since language revitalization must be grounded in the participation of community members, academic partners in language revitalization serve as external resources providing input in the process, typically in support roles that may involve curriculum or materials development, grant writing, program administration, or training. Such activities, when successful, are therefore enabling of the empowerment process of language revitalization.

In sum, while models of linguistic fieldwork leading to language documentation span the entire continuum of ethical research practices and a range of relationships between academic and community participants in the research, language revitalization processes, providing they are successful, are inherently empowering and academics play support roles that enable these empowerment processes. We now begin to examine these issues with regards to training.

3. Training and empowerment

There is a natural connection between the concepts of training and empowerment, since imparting new skills and knowledge in principle allows for the development of capacity and autonomy, both aspects of community empowerment. Training can be delivered in a variety of contexts and locations. An important difference is between programs that are *in situ*, focused on a single language community in the context of a revitalization or maintenance project, and programs that are *ex situ*, with participants traveling regionally, nationally, or internationally to colleges, universities, museums, archives, or other institutions for training programs designed to address participants from multiple language communities. There are advantages and disadvantages of both types of programs. While *in situ* training can involve a large number of participants in a single language revitalization

project, typically only a small number of members from any given community participate in *ex situ* programs. Such people might return to their home communities and pass on what they've learned to other members, but the extent to which this is done, and the effectiveness of the transmission, varies. On the other hand, *ex situ* programs, especially those that bring participants from many endangered-language communities, provide an important perspective on the broader context of language endangerment, allow for the formation of supportive networks among language activists, and can provide a broader range of skills and expertise. Both types of training are beneficial and complementary.

The Institute on Field Linguistics and Language Documentation (InField, referred to as the Institute for Collaborative Language Research (CoLang) in 2012) is an example of an *ex situ* program, as it pulls participants from across the globe to a university setting in the United States.[3] The magnet approach not only underscores the point that language endangerment is an international phenomenon; it also provides participants with a broad variety of perspectives on issues surrounding language endangerment, including the academic-community relationship. InField is especially interesting in this regard as it is designed to bring together both members of endangered-language communities and academics for shared teaching and learning.

An obvious motivation for designing InField so that it served a 'blended' audience was a clear need for training for all parties involved in language documentation and conservation. This was especially evident with regards to the rapidly changing technological environment around the turn of the 21st century. In retrospect, it also seems that the idea of bridging these academic-community constituencies involved in LDC was inspired by the literature on ethics in fieldwork discussed above, although at the time this link was not explicit. However, both directors were aware of the frequently fraught relationship between linguists and members of indigenous communities and of the need for greater understanding between academic and community partners, especially in the North American context. But perhaps the strongest motivation for the broad participant base was the conviction that everyone involved in language documentation and conservation has both skills to offer and much to learn, and that people are more effective working together towards shared understanding and shared goals than when they are working separately.

InField was organized in two parts: the first two weeks of the institute comprised workshops on topics in language documentation, revitalization, and

---

3. InField 2008 was funded by a grant from the National Science Foundation Documenting Endangered Languages Program (BCS-0724221).

maintenance; this was followed by four weeks of three concurrent intensive field-training courses, each offered on a different language with a different instructor and language consultants. The workshops were chosen through an open call for workshop proposals, which was widely distributed through mailing lists. The response to the call for proposals was strong, with more than 40 submissions from scholars and community language workers from around the globe. The diverse set of proposals covered core areas of LDC such as technologies, theories and models of documentation and revitalization, data management, ethical and social aspects of endangered language work, resource creation, and skills and methods such as grant writing, orthography development, and 'best practices' for language documentation. The proposals were circulated to an international organizing committee, which undertook the task of choosing among them and crafting a comprehensive and coherent curriculum. The decisions were based both on the content of proposed workshops and on the qualifications of the instructors.

The instructors, selected based on their expertise and direct connections to community-based language projects, were drawn from all over the globe. Community-member language activists were prioritized for selection as instructors. However, only a handful of proposals were submitted by community members; and out of the twenty-seven instructors that taught at the 2008 InField, only three were members of endangered-language communities. In several cases people had submitted similar or complementary proposals, so the organizers asked many instructors to co-teach their workshops. This was a challenge for instructors who were not previously acquainted with each other or had to coordinate planning at a distance, but was also highly rewarding for many instructors and students and made an important contribution to the collaborative culture of the institute.

3.1 The InField workshops

Returning to the theme of empowerment, the workshops didn't constitute empowerment *research*, as they weren't research-driven, although they were designed to enhance and enable the research projects of participants. Being *ex situ*, and having a geographically and culturally diverse range of participants, they did not draw on local traditions or involve community members in negotiating goals or outcomes. On the other hand, some dimensions of the workshops can be said to have been empowering. They were clearly *capacity-building*, providing skills and knowledge that directly enhanced and enabled community initiatives. For example, Kennedy Bosire and Gladys Machogu directly applied recording and data management skills from the workshops in adding multimedia files to a Lexique Pro version of their Ekegusii dictionary, which they put online in 2010. Another example is

linguist and community activist Susan Gehr, who applied her newfound Toolbox skills to advance the Karuk community's language revitalization work. In addition, some workshops were *participatory and cooperative*. Many of the technology workshops had hands-on portions. Other workshops had students working in small groups, purposely pairing academics with community-based activists to provide input on each other's projects. And although the workshops necessarily had the instructor-student divide, the instructor teams who co-taught the workshops served as models of collaborative interaction. Finally, a significant concept underlying the InField ethos is that everyone has something to share and much to learn. This view that all participants are teachers and all are learners was stated explicitly at the opening of the institute and at other times; it exerted a leveling effect that promoted mutual respect and the concept of partnership.

That the workshops were viewed as empowering of community members can be seen in some of the comments provided in written evaluations collected at the end of the institute or in post-institute interviews. (All quoted comments are taken from Siemens 2010.) The observation that the workshops are empowering community members to work independently in line with their own goals is recorded in the following:

> Greater community-academy collaboration (evidence of which we saw at Infield) demonstrates that documentation expertise is being transferred into community hands, and is therefore being applied by speakers (& aspiring speakers) in ways that suit community-designated interests. Very inspiring!

The following comment, from a community-based participant, suggests that the expertise has been fully shifted from academic to community members, with community-based language activists as experts who can serve as instructors at future institutes: "From here as community speakers, we are equipped to be instructors in the forthcoming InFields. Use your products!"

In addition to providing a degree of empowerment for community-based participants, many of the workshops also modeled explicitly collaborative interactions that impacted the academic participants. Again the comments collected in evaluations and interviews are instructive:

> I think [what affected me most] was the different kinds of models that were presented about language documentation and linguistic work, and linguistic scholarly goals in relation to community goals and ways in which those can be aligned and ways in which sometimes they are not aligned. So there were moments in different workshops where you can kind of see academic goals diverging from community goals and vice versa, or ways in which that's happened in the past. I think that what I learned from that for my own work was that it is really possible to bring those things in line if you decide ahead of time that that's how you want to work.

> You have to understand the community's view of you and shift your goals to match their goals. You can still satisfy the scientific hunger because your work is now informed by the people and culture – it's more meaningful.
>
> There are different definitions of collaboration and some definitions of collaboration are more collaborative than others. Some ways of working collaboratively are more thoroughly shared – a shared process – than others. That was really interesting and helped a lot when it came time for me to figure out how I was going to create a collaborative project and build that relationship.

These comments demonstrate a conceptual shift that occurred for some of the academic participants towards views of fieldwork that are likely to be more collaborative and empowering of the communities in which they work. The workshops can thus be seen as promoting empowerment models of research and thus indirectly supporting community empowerment.

3.2 Field training

Following the workshops, 30 of the participants continued their training in one of three field-training classes. Each class consisted of one or two native speaker teachers, one linguist-instructor, and about 10 students. The classes were intensive: they met daily for three hours and students also met in pairs with a speaker for two hours every other day. Each of the classes was unique in its goals, style, the situation of the language being studied, and the backgrounds of the participants. Each reflected different approaches to collaboration and the balance between the needs of the academic community and the needs of the language community. This section presents a brief case-study of each of the classes, based on the evaluations, interviews, and first-hand observations.

The Mende class was the most like a traditional university field-methods course. Mende, with over 1.4 million native speakers, is a minority language, but not an endangered language. There was one native speaker consultant, Taziff Koroma, who is a professor of Mende in his home country, Sierra-Leone. Thus, his knowledge of linguistic analysis was sophisticated, as were his ideas about pedagogy. Tucker Childs, a professor at Portland State University, was the linguist-instructor for the class, and he and Taziff already had a long established working relationship. The students in the class were linguistics graduate students. The class was structured in such a way that students were expected to consult existing resources on the language, related languages, and geographically close languages in order to get as far as possible in the analysis of the language structures. There was also some lecture time devoted to teaching relevant background information and learning the mechanics of computer software for linguistic analysis. Each student chose an aspect of the language structure to focus on, and all contributed

to a class wiki with the goal of collecting examples and analyses that would form a grammatical sketch. The focus of the class was on honing research and analytical skills, and how to work with a native speaker to create a linguistic sketch of the language.

The Ekegusii class included two native speaker consultants, Gladys Machogu and Kennedy Bosire. The two came to InField as language activists; they had already begun the Ekegusii Encyclopedic Projects – a project to document the Ekegusii language – before learning about InField. Ekegusii is an endangered language of Kenya. While it still has over 2 million speakers, its use is stigmatized and restricted in urban communities, and many young children are not actively learning the language. The linguist-instructor for the class was Carol Genetti, a professor at UC Santa Barbara. The students in this class were linguistics graduate and undergraduate students and one interested non-linguist.

This class aimed to replicate the fieldwork experience for students. They were given very little in the way of background information on the language and worked intensively as a group to discover the structures of the language from the sounds to the grammar. The students in this class also contributed to a class wiki and worked on an analytical presentation of some aspect of the language. However, as the class progressed, the focus of the class shifted from practicing the methods and tools learned during the two weeks of workshops to actively collaborating on the real and ongoing project of the consultants. This shift was unintentional and natural, given that Kennedy and Gladys had traveled from Kenya for the purpose of learning skills needed for their project. The class was able to help them in the expansion of their interactive dictionary, adoption of an orthography that included tone marking, understanding of lexical and grammatical categories, and work towards a collection of folk tales. By the time the institute came to an end, the consultants had learned several key skills for their documentation project, had greatly deepened their understanding of Ekegusii grammar, and had made valuable connections with students and linguists that would lead to further collaboration.

The Kwak'wala class was quite special in the make-up of its participants. There were two native speaker consultants, elders Beverly Lagis and Daisy Sewid-Smith. In addition, there were four Kwakwaka'wakw community members in the class, two students who were involved with Kwak'wala language projects, linguistics graduate students, and two activists involved in language revitalization with another community. Patricia Shaw, a professor at the University of British Columbia, was the linguist-instructor for the class.

Kwak'wala is a highly endangered language, and this fact shaped many aspects of the course. The elder speakers that worked with the class were from different dialect and community groups, each with their own existing orthographic traditions as well as ongoing language revitalization projects. The younger community

members in the class had varying degrees of proficiency in the language, and varying degrees of exposure to linguistics and to the higher education system. The graduate students in the class also came into the class with different levels of exposure to the language and to the community of speakers.

Because of the diverse backgrounds of participants and the small number of fluent native speakers among the communities, in this class much attention was explicitly given to working out plans for meeting community members' goals and working together in ways that were appropriate for all participants. Thus, although it was *ex situ*, this class most strongly modeled the participatory action-based research frameworks being proposed in the literature on fieldwork. Issues surrounding language access, intellectual property rights, and collaboration were primary, and only after such issues had been thoroughly addressed, did the technical work of documentation proceed. Thus, the Kwak'wala field training embodied what InField had set out to do – to bring different people together for mutual learning, teaching, and working toward shared goals and understanding. Based on the continuation of relationships and work begun during Kwak'wala field training, the time and effort dedicated to working on and through ethical concerns was well worth it in leading to productive collaboration and community-based language projects.

Of the three classes, both the Ekegusii class and the Kwak'wala class were directly empowering of their community participants. In both cases, the participants returned to their home communities to continue with productive work. In the Ekegusii case, this led to the completion of the dictionary (Bosire & Machogu 2013) and the expansion of the documentation project. Kennedy Bosire has served as an instructor at both the 2010 InField and the 2012 CoLang institutes, where he has continued his training and his collaboration with linguist Carlos Nash, who was a graduate student participating in the Ekegusii class in 2008. In the Kwak'wala class, all of the community member students have continued their active work on the language. They have been taking part in a wide range of projects, from teaching community-based language classes, to language documentation, to developing resources for children. Of these, Mike Willie served as an instructor at the 2010 InField. Two graduate students from InField 2008 have continued active collaboration with the Kwakwaka'wakw community. Co-authored talks that presented this work and discussed the process of negotiating difference within the context of the work were presented at the 2009 and 2011 International Conference on Language Documentation and Conservation (Cranmer et al. 2009; Rosenblum, Cadwallader, Nicholson & Willie 2011). This course was also clearly empowering for the community participants.

In sum, although the field training portion of InField was a training project, and community external, because the classes were research-based and grounded

in the languages, they were still empowering of the community-based participants from endangered-language communities. They were participatory, and directly addressed the needs of the participants, providing valuable information that furthered their work within their communities. The community members and researchers were viewed as jointly engaged contributors, working together in partnership. The goals and outcomes of the research arose organically in the case of the Ekegusii class and were determined through explicit discussion and negotiation within the Kwak'wala class. And both classes were capacity building, enhancing work in community initiatives. The following quotes, taken from evaluations at the end of the field training class, give voice to the sense of empowerment gained by community participants:

> I did not consider myself a linguist when I arrived, but I now feel as though I am in a position to contribute to the documentation aspect of language revitalization in a concerted and effective way. I learned so many terms, skills, and strategies that I can apply to my work.
>
> As a non-linguist, but as language activist I found InField training to be an enriching experience. The forum, format, content, and sequencing of the components contributed to moments of epiphany in connection with my own language. I am inspired to continue with intensive future work.
>
> I gained an enormous amount of skills through workshops and through networking with the participants. I had very little linguistic training prior to coming here, but I came here because language revitalization is my life. Thank you for this tremendous opportunity. "InField" has confirmed my ambition to pursue an education in linguistics. I will be recommending InField to my community members and others. Thank you.

These comments demonstrate that the sense of empowerment came from developing a foundation in their language work and the application of those skills through the field training class. The course helped the students to develop both the skills and the confidence to continue independently, and this in turn resulted in the higher levels of commitment to the work revealed here.

3.3 Further dimensions of empowerment

We return now to the situation of InField in a community-external environment and its impact on participants. In bringing together people from a geographically diverse range of endangered-language communities for shared teaching and learning, InField as a whole provided a dimension of empowerment not attainable in community-based settings. Instead of local grounding, the institute grounded the work of the participants within an international social movement,

hence provided a broader framing of the local work and its social impact. The group experience provided acknowledgement, encouragement, and direct support of the community-based participants, thus strengthening their sense of purpose and confidence as they returned to their home communities. In addition, for some participants their involvement in an international institute was seen as prestigious at the local level; it validated their work and strengthened their status (i.e. increased their social power) within their communities. Finally, the institute provided multiple role models of community language activists as agents of social change, making an impact not only within but also beyond their local communities. A number of people expressed an interest in participating as instructors in future iterations of the institute, hence recognized the value of their own skills and scope of potential social impact. We can thus see that *ex situ* training programs promote different dimensions of empowerment than programs grounded in the home community.

4. Conclusions

As discussed above, the ethics literature distinguishes research paradigms by the nature of the social relationships between community and academic partners and by the locus of control over research goals, methods, and outcomes. We additionally complicate the discussion by adding in the extent to which the research is likely to constitute social action, in the senses discussed above.[4]

In ethical research the power is on the side of the academic. The primary goal is the discovery of knowledge and there is no explicit intention to create social change or to contribute to the educational or social advancement of the 'subjects'. Advocacy research is intended to bring about social change with the academic partner as advocate, who still retains primary control over the research process. Advocacy research does not necessarily involve community engagement in the research per se, so does not necessarily constitute a social-action process in the senses discussed above. Empowerment research does bring in the component of community engagement, although control over the project might still be on the side of the researcher. This shifts in CBLR, where community members are additionally responsible for the goals and outcomes of the research, so are the most strongly engaged in ensuring that the research is designed and carried out for the betterment of the broader community. Training can happen at any of these levels

4. We refer here to abstract prototypical exemplars of research types. In reality, each research situation is unique and profoundly shaped by the individuals involved, their views and ideologies, community structures, the broader linguistic ecology, and a host of other factors.

(e.g. a linguistic consultant can be trained to make recordings and transcribe in support of an ethical research project), but is likely to be both broader and deeper in empowerment research and CBLR.

Training programs constitute a distinct type of activity, with an explicit goal of capacity development. In empowering social action, training programs directly support language goals of communities undertaking conservation work. They thus significantly alter the academic-community relationship, with the academic/trainer in a support role as community resource. Training programs vary in the extent to which they incorporate research (parallel to the ways in which research paradigms vary in the extent to which they incorporate training). Training programs can be organized into a separate typology which is independent from, but overlaps with, the typology of research paradigms. *Ex situ* training programs, such as InField, promote empowerment dimensions that are different from those afforded by community-based programs, so have a unique position within the typology. Worldwide, training programs vary in numerous ways, suggesting a complex typology for future study.

References

Beier, Christine. 2009. Implementing community involvement in collaborative language documentation: The Iquito case. Talk presented at the 1st International Conference on Language Documentation and Conservation, University of Hawai'i at Manoa, March.

Benedicto, Elena; Antonlín, Demetrio; Delores, Modesta; Fendly, Gloria; Gómez, Tomasa; Salomón, Elizabeth; Viñas de Puig, Ricard & Eggleston, Alyson. 2009. Participatory documentation: The Mayangna linguists team of Nicaragua. Talk presented at the 1st International Conference on Language Documentation and Conservation, University of Hawai'i at Manoa. March.

Bosire, Kennedy Momanyi & Gladys Kwamboka Machogu. 2013. *Ekegusii-English Dictionary with English Index*. The Republic of Kenya: Ekegusii Encyclopedia Project.

Bowern, Claire. 2008. Consultants and field locations. *Linguistic Fieldwork: A Practical Guide*. Houndmills: Palgrave Macmillan.

Cameron, Deborah, Frazer, Elizabeth, Harvey, Penelope, Rampton, M.B.H. & Richardson, Kay. 1992. *Researching Language: Issues of Power and Method*. London: Routledge.

Cameron, Deborah, Frazer, Elizabeth, Harvey, Penelope, Rampton, M.B.H. & Richardson, Kay. 1993. Ethics, advocacy and empowerment: Issues of method in researching language. *Language and Communication* 13(2): 81–94.

Chelliah, Shobhana & DeReuse, Willem. 2010. *Handbook of Descriptive Linguistic Fieldwork*. London: Springer.

Craig, Colette. 1993. Comment. *Language and Communication* 13(2): 97–100.

Cranmer, Laura, Smith, Julianne & Shaw, Patricia. 2009. Reconciling difference and building trust: International collaboration in indigenous language revitalization. Talk presented at the 1st International Conference on Language Documentation and Conservation. University of Hawai'i at Manoa, March.

Crowley, Terry. 2007. *Field Linguistics: A Beginner's Guide.* Oxford: OUP.

Czaykowska-Higgins, Ewa. 2009. Research models, community engagement, and linguistic fieldwork: Reflections on working within Canadian indigenous communities. *Language Documentation and Conservation* 3(1): 15–50.

Dimmendal, Gerrit. 2001. Places and people: Field sites and informants. In *Linguistic Fieldwork*, Paul Newman & Martha Ratliff (eds.), 55–75. Cambridge: CUP.

Dobrin, Lise. 2008. From linguistic elicitation to eliciting the linguist: Lessons in community empowerment from Melanesia. *Language* 84(2): 300–324.

Dwyer, Arienne. 2006. Ethics and practicalities of cooperative fieldwork and analysis. In Gippert, Himmelmann & Mosel (eds), 31–66.

Furbee, N. Louanna & Stanley, Lori A. 2002. A collaborative model for preparing indigenous curators of a heritage language. *International Journal of the Sociology of Language* 154: 113–128.

Grinevald, Colette. 2003. Speakers and documentation of endangered languages. In *Language Documentation and Description,* Vol. 1, Peter K. Austin (ed.), 53–72. London: The Hans Rausing Endangered Languages Project, School of Oriental and African Studies.

Gippert, Jost, Himmelmann, Jost Nikolaus & Mosel Ulrike (eds). 2006. *Essentials of Language Documentation.* Berlin: Mouton de Gruyter.

Guérin, Valérie & Lacrampe, Sébastien. 2010. Trust me, I am a linguist! Building partnership in the field. *Language Documentation and Conservation* 4: 22–33.

Harmon, David & Loh, Jonathan. 2010. The Index of Linguistic Diversity: A new quantitative measure of trends in the status of the world's languages. *Language Documentation and Conservation* 4: 97–151.

Himmelmann, Nikolaus. 1998. Documentary and descriptive linguistics. *Linguistics* 36: 161–195.

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

Hinton, Leanne. 2011. Revitalization of endangered languages. In *The Cambridge Handbook of Endangered Languages,* Peter K. Austin & Julia Sallabank (eds), 291–311. Cambridge: CUP.

Holton, Gary. 2009. Relatively ethical: A comparison of linguistic research paradigms in Alaska and Indonesia. *Language Documentation and Conservation* 3(2): 161–175.

Horvath, Peter. 1999. The organization of social action. *Canadian Psychology/Psychologie canadienne* 40(3): 221–231.

Jukes, Anthony. 2011. Researcher training and capacity development in language documentation. In *The Cambridge Handbook of Endangered Languages,* Peter K. Austin & Julia Sallabank (eds), 423–446. Cambridge: CUP.

Leonard, Wesley & Haynes, Erin. 2010. Making "collaboration" collaborative: An examination of perspectives that frame field research. *Language Documentation and Conservation* 4: 268–293.

Michael, Lev. 2009. Developing infrastructure for team-based research: The module-and-seminar model. Talk presented at the 1st International Conference on Language Documentation and Conservation. University of Hawai'i at Manoa, March.

Minkler, Meredith & Wallerstein, Nina. 2008. Introduction to CBPR: New issues and emphases. In *Community-Based Participatory Research for Health: From Process to Outcome,* 2nd edn, Meredith Minkler & Nina Wallerstein (eds), 5–24. San Francisco CA: John Wiley and Sons.

Mosel, Ulrike. 2006. Fieldwork and community language work. In Gippert, Himmelmann & Mosel (eds), 67–86.

Nevins, Marybeth E. 2004. Learning to listen: Confronting two meanings of 'language loss' in the contemporary White Mountain speech community. *Journal of Linguistic Anthropology* 14(2): 69–88.

Penfield, Susan D.; Serratos, Angelina; Tucker, Benjamin V.; Flores, Amelia; Harper, Gilford; Hill Jr., Johnny & Vasquez, Nora. 2008. Community collaborations: Best practices for North American indigenous language documentation. *International Journal of the Sociology of Language* 191: 187–202.

Rice, Keren. 2006. Ethical issues in linguistic fieldwork: An overview. *Journal of Academic Ethics* 4: 123–155.

Rice, Keren. 2009. Must there be two solitudes? Language activists and linguists working together. In *Indigenous Language Revitalization: Encouragement, Guidance, and Lessons Learned*, Jon Reyhner & Louise Lockhard (eds), 30–52. Flagstaff AZ: Northern Arizona University.

Rice, Keren. 2010. The linguist's responsibilities to the community of speakers: Community-based research. In *Language Documentation: Practice and Values,* Lenore A. Grenoble & N. Louanna Furbee (eds), 25–36. Amsterdam: John Benjamins.

Rice, Keren. 2011a. Ethics in fieldwork. In *Documenting Endangered Languages: A Fieldworker's Companion*, Nicholas Thieberger (ed.). Oxford: OUP.

Rice, Keren. 2011b. Documentary linguistics and community relations. *Language Documentation and Conservation* 5: 187–207.

Rosenblum, Daisy; Cadwallader, Elizabeth; Nicholson, Deanna & Willie, Mikael. 2011. Collaborative approaches to transcription for community-based documentation of spontaneous interactive speech in Kwak'wala. Talk presented at the 2nd International Conference on Language Documentation and Conservation. University of Hawai'i at Manoa, February.

Tsey, Komla. 2010. Making social science matter? Case studies in community development and empowerment education research in rural Ghana and Aboriginal Australia. *Asian Social Science* 6(1): 1–11.

Tsunoda, Taskau. 2006. *Language Endangerment and Language Revitalization*. Berlin: Mouton de Gruyter.

Wallerstein, Nina. 1992. Powerlessness, empowerment, and health: Implications for health promotion programs. *American Journal of Health Promotion* 6(3): 197–205.

Weber, Max. 1978. *Economy and Society*. Berkeley CA: University of California Press.

Wilkins, David P. 1992. Linguistic research under Aboriginal control: A personal account of fieldwork in Central Australia. *Australian Journal of Linguistics* 12: 171–200.

Woodbury, Tony. 2003. Defining documentary linguistics. In *Language Documentation and Description,* Vol. 1, Peter K. Austin (ed.) 35–51. London: The Hans Rausing Endangered Languages Project, School of Oriental and African Studies.

Woodbury, Tony. 2011. Language documentation. In Peter Austin and Julia Sallabank (eds.) *The Cambridge Handbook of Endangered Languages*. 159–186. Cambridge: Cambridge University Press.

Yamada, Racquel-Maria. 2007. Collaborative linguistic fieldwork: Practical application of the empowerment model. *Language Documentation and Conservation* 1(2): 257–282.

How to avoid pitfalls in documenting endangered languages

Sarah G. Thomason
University of Michigan

Conducting fieldwork on an endangered language resembles fieldwork projects on non-endangered languages in most respects, but there also some differences that require extra attention when one studies an endangered language in the field. This paper concentrates on the differences. It covers such topics as getting access to fluent native speakers and community approval for the research; data-collection techniques (a variety will be needed); individual variation, which can be much more problematic in an endangered language than in other languages; and dictionary-making procedures, which may require consultants to dig deep in their memories to come up with words they haven't spoken or heard for many years.

1. Introduction

This is a 'how-to' paper with a target audience of linguists who are about to undertake their first major fieldwork project on an endangered language. My focus is on primary documentation – that is, starting from scratch with fieldwork on a previously undocumented or underdocumented language, specifically an endangered language. The paper is not meant as a general introduction to fieldwork; for that, readers should turn to one or more of the increasingly numerous books on the topic, e.g. Sakel & Everett 2012; Bowern 2008; Crowley 2007; Chelliah & de Reuse 2010; Vaux et al. 2007; Thieberger (ed.) 2012; the pioneering book by Samarin (1967); or the excellent volume edited by Newman and Ratliff (2001).

Some of the issues that arise in documenting endangered languages are the same ones that come up with any primary documentation project: making initial contacts with the community, getting permission to conduct fieldwork (both from a university's Institutional Review Board and from the speech community), organizing a field session, preparing carefully for each session, selecting techniques for collecting lexical and grammatical data, trying hard not to offend anyone (e.g. by trying to elicit taboo words and concepts), and so forth. But there are also special circumstances that make fieldwork on an endangered language especially

challenging. Among these circumstances are the need to start by approaching community leaders rather than potential language consultants; the need for extra flexibility in organizing the work, given the (probable) advanced age and precarious state of health of the consultants; the need to use a variety of methods in collecting data from elderly speakers who might not have used their language regularly for decades; the need to accept most of the variant forms provided by different speakers, even when the speakers disagree among themselves about the acceptability of some variants; and the need to ensure that the results of the research are usable by the community as well as by linguists. In addition, in some cases – both with research on nonstandard dialects and with research on languages that haven't (yet) been standardized – processes of standardization complicate speakers' assessment of certain forms that they and other community members produce, and the linguist must deal with the effects of varying grammaticality judgments that arise from this cause.

Most of my examples will be drawn from my own fieldwork in two very different situations: documentation of word-formation patterns in endangered nonstandard dialects of the language formerly known as SerboCroatian; and, more extensively, primary documentation of the Salish-Pend d'Oreille (Montana Salish) language that is spoken on the Flathead Reservation in northwestern Montana. The youngest speakers I worked with in the former Yugoslavia in 1965–1966 and 1967 were in their 60s, had never attended school, and were not all literate; I traveled around the country to selected villages, where the dialects were even then under heavy pressure from Standard SerboCroatian. The youngest speakers I've worked with in Montana, starting in 1980, are (or were) over 70, are fully fluent in English as well as in their native language, and are fully literate in English. It has been two or more generations since any children were raised mainly speaking Salish-Pend d'Oreille, in part because many of the current elders are survivors of school systems (religious boarding schools and also public schools) in which they were beaten for speaking their language. The language as spoken by the few remaining traditional native speakers will therefore be gone within the next 20–30 years.

The rest of the paper is organized into five main sections: beginning the fieldwork (§2), managing field sessions (§3), dealing with the thorny problem of individual variation (§4), and an extended example – dictionary-making (§5). Section 6, finally, is a brief conclusion. Knowledgeable readers will notice significant gaps in the coverage here of topics that are relevant to a fieldwork project; I do not, for instance, discuss technological issues of data management. To fill such gaps, readers should consult book-length discussions of fieldwork such as the ones listed above, as well as individual articles on topics missing here (e.g. Margetts & Margetts 2012; Thieberger & Berez 2012).

2. First steps

The first goal, when starting a documentation project on an endangered language, is to get access to native speakers of the language. Access to the people you need to work with is not usually a problem with a non-endangered language: the language communities are usually bigger, so even if some community members might object to your being there, others are likely to be willing to work with you and relatively unlikely to be subject to community sanctions for doing so. The situation with an endangered language is often (though not always) quite different. When a community's culture is under threat because of assimilation to a dominant group's culture and concomitant shift to their language, the community leaders will often wish to keep what they have left of their culture by excluding outsiders – especially outsiders who are perceived to be members of, or in cahoots with, the dominant group. Many fieldworkers have encountered community suspicions that linguists (and anthropologists) are trying to steal their language and use it to get rich.

The widespread belief that linguists and anthropologists are the bad guys makes it absolutely vital to treat members of the language community with respect and even deference. In practice, this means explaining carefully what work you want to do and why, how the results of the research will be used, and – most importantly – how the research will benefit the community. Giving back to the community in which you conduct fieldwork is now an imperative in most or all field situations, again especially in the case of endangered languages.

You should get approval for your work from community leaders before approaching the elders who are fluent speakers, the keepers of the language. Approval mechanisms range from an informal verbal O.K. all the way to a formal written protocol. If the community asks for guarantees embodied in a written protocol, follow their wishes. Such protocols will lay out the conditions of your work (which elders you can work with, payment expected, etc.) and will also specify what you can and can't do with the data you collect. When I worked in the old Yugoslavia, my first stop when I arrived in a village was the office of the local Communist official, to explain myself and my work and make sure he (it was always a man) had no objection to my being there; once he said yes, I could approach any member of the community. In the case of an indigenous language in the United States or Canada, the prospective fieldworker should approach the appropriate authority, typically the tribal chief, the tribal council, or the council of elders. My work in Montana began when the Flathead Culture Committee (as it was then called; it is now the Salish-Pend d'Oreille Culture Committee) invited me to conduct a workshop to help Culture Committee members and tribal elders learn how to read and write their new IPA-based alphabet. All my fieldwork since then has been under the auspices of the Culture Committee. In other parts of the world

there may be different avenues to approval of your research, and of course in some instances no approval will be necessary. But the fieldworker always needs to check carefully to make sure s/he avoids offending the people whose good will is necessary if the fieldwork is to take place. Other ethical and even legal issues also arise in fieldwork contexts; see, for instance, Rice 2012 and Newman 2012 for discussion of some of the most important considerations in these areas.

Fieldworkers who are connected with a U.S. university must also follow their university's rules about getting approval from the Institutional Review Board (IRB), because the IRB will view your fieldwork as human subject research. In some parts of the world this restriction makes excellent sense – speech can kill, and your highest priority should be to avoid causing harm to your language consultants. In other parts of the world, calling your consultants 'human subjects' feels like an offensive denial of their expert status. But that feeling will not exempt you from the necessity of having IRB approval for your fieldwork.

Once you meet the people you want to work with – perhaps the last fluent speakers of the language – your preeminent task is to get along with them. Never argue with them, never disagree, always show respect for their opinions as well as for their knowledge of the language you want to document; and be sensitive to local cultural norms. In Yugoslavia, I found that I could not work with men in the villages I visited: they were happy to talk to me at great length, but they wouldn't answer my questions, because it was culturally inappropriate for a woman to question a man. Women would answer my questions, but they wouldn't talk at all when their menfolk were present; so I had to find ways, often devious ways, of getting the women alone so that I could elicit data systematically.

Another pitfall there was the occasionally tense political situation. In a village near the Albanian border, for instance, a woman I was trying to interview thought I was from the secret police, and she declined to say anything negative about anybody – a serious problem for me, because among other things I was looking for pejorative terms. ("How do you say, about someone you dislike, 'He has a big nose?'" "God gave him a big nose; it's not for me to criticize.") I was also certain that the Communist authorities would be unhappy if they knew about the people who wanted to tell me how dreadful Communism was, so I always tried to cut those conversations short. The best strategy is to avoid political discussions entirely when you're in a new cultural context.

If you offend your consultants, they won't want to work with you. You may develop close working relationships and even personal friendships with them, but you should always be aware of one salient fact: you need them more than they need you. If their language is gravely endangered, you cannot risk doing anything that might make them skip scheduled working sessions with you or refuse altogether to work with you. There might be no other fluent speakers available, especially as any offense you give will probably become known to the whole community.

It may seem obvious that you should make every effort to learn to speak the language – that the more fluent you become, the better. But with an endangered language this goal might be unrealistic, and it may also be an inappropriate goal for cultural reasons. It's easy to see why it might be unrealistic to hope to achieve fluency: if there are only a few remaining fluent speakers, and if they are all elderly, they're very likely bilingual in the dominant language now spoken by the rest of the community; it may now be their dominant language. If they have fallen out of the habit of speaking their native language regularly, speaking their language to you – an outsider – might feel like too much effort, especially when they speak to younger relatives and other community members in the dominant language.

It can be harder to detect cultural reasons that make it inadvisable to try to become fluent in the language you are investigating. There are reports from various parts of the world of speakers deliberately withholding their language from outsiders (see Thomason 2001: 161–162, 191 for discussion and references). Communities whose languages are endangered, as noted above, are perhaps more likely than other speech communities to be reluctant to have outsiders speak their language. And if your target language is a nonstandard dialect, speaking the dialect will almost certainly be counterproductive. When I conducted fieldwork in the old Yugoslavia, I spoke only Standard SerboCroatian, but I was looking for word-formation patterns in nonstandard dialects. Anyone with significant exposure to the standard dialect would try to speak it to me, which was unfortunate; that's why I sought out consultants who had never been to school and whose control of the standard dialect could reasonably be assumed to be minimal. (In the mid-1960s, the villages I visited had few or no television sets, though some people did have radios.) If I had tried to speak the local dialect in a village, it would have sounded as if I were making fun of the villagers, which would have been even more unfortunate. My solution – consultants old enough to have had no formal schooling but young enough to be sharp-witted still – worked quite well.

In Montana, the few remaining fluent elders now speak their language only rarely, primarily when they meet in a group with me in the summer and during periodic elders' meetings the rest of the year. They don't expect anyone else, either tribal members or outsiders, to speak it at all. I've found that it makes them a bit nervous when I laugh at the jokes they tell each other, so I don't usually let them know when I understand their side conversations during our sessions together. It would be intrusive for me to try to go further toward fluency. I've heard comparable stories from other fieldworkers. One told me about working with an indigenous community in Mexico whose members trusted him and were happy to work with him. He worked hard to learn to speak the language. Eventually he understood what they were saying in it, whereupon they told him, "Now we don't trust you anymore." So don't assume that you should learn to speak the target language fluently; your consultants might not appreciate it. (My experiences in this matter

are not universal, of course. For the other side of the picture, see Everett 2001: 167 and Dimmendaal 2001: 72–73.)

The final topic in this section concerns the selection of consultants. Some of the recent textbooks on field linguistics suggest criteria for choosing speakers to work with on documentation projects. Sakel & Everett, for instance, say that a good consultant will be a good story-teller, friendly, a speaker of the correct dialect, a native speaker of the language, without any speech impediment, able to reflect on the language as a formal system, and well-respected in the community (or at least not marginalized). But if you are investigating a gravely endangered language, you almost certainly won't have the luxury of finding and selecting the ideal consultant; if the language has only a few fluent speakers left, you will work with whoever is available and willing. Moreover, there is (of course) no such thing as the ideal consultant: everyone has her/his idiosyncrasies, and the wise field linguist works with, or if necessary around, the consultant's quirks.

The main criteria for selecting consultants are these: native-speaker fluency and willingness to serve as a consultant for your project. You will find that different speakers have different skills. For instance, one of the Salish-Pend d'Oreille elders that I work with has beautifully clear pharyngeal consonants; but his ejectives are so lenis that I can hardly hear them. Another example: two elders I worked with over twenty years ago were wonderfully patient about going over old tape-recorded stories sentence by sentence so that I could transcribe them; but they wouldn't tell me stories of their own. Another elder who occasionally drops in on our sessions has a talent for fine-grained semantic explanation of subtle differences between apparent synonyms; but he very rarely volunteers any forms when I'm eliciting data. Still another elder was a workhorse at the age of 90, willing – even eager – and able to work long hours without a break; but he dominated our sessions so much that the other elders could contribute only rarely.

In sum, with a primary documentation field project, you play to your consultants' strengths: elicit stories from story-tellers, go over old tape recordings with patient consultants, encourage an especially talented elder to provide a context where a word or phrase might be used, and so forth. Don't hold out for the ideal language consultant; you can get a great deal of data from a consultant who has flaws as a language worker.

3. Managing field sessions

Three points should be kept firmly in mind when you plan your field sessions: the work should be paced carefully during the day; thorough preparation for each session is vital; and you will need to suppress your own time anxiety. When planning a day's work in the field, you will also want to use a variety of

data-gathering techniques, both to avoid boring your consultants and because no single technique will give you all the kinds of data you'll need for a comprehensive description of the language. In this section I'll discuss each of these points in turn.

When you are investigating an endangered language, there is a high probability that all your consultants will be elderly, which means that almost all of them will have limited reserves of energy. It may also mean, as noted above, that decades-long replacement of their native language by the dominant language in everyday conversation has left them rusty: they can carry on casual conversations easily, but they often have to struggle to remember words for many items and concepts, and translating sentences into their language also requires extra effort. Concentrating hard over several hours is tiring, even for younger people. (That's why, for instance, world-class chess players need to be in excellent physical condition.) Two ways of protecting your consultants from exhaustion are to work with a group of elders rather than just one consultant for an entire day, and to allow for brief periods of relaxation during the day – at least a mid-morning break, a lunch break, and a mid-afternoon break.

During my weekly visits to the Flathead Reservation in the summer, there are also brief periods of casual conversation among the elders as the work proceeds: particular words or topics trigger memories that the elders comment on, or a visitor wanders into the room with a question, or someone mentions a current news item. These interruptions contribute to a relaxed atmosphere around the table, and they make it easier for everyone to concentrate again after the interruption is over. If you set a leisurely pace, you will also avoid wearing yourself out, and you'll be less likely to miss vital points. Our working day officially begins at 9:00 and ends at 4:00; in practice, there are perhaps four or five hours of concentrated language work. I don't know how tired the elders are by 4:00, but I know that I am always exhausted. Nowadays fieldworkers routinely tape-record and/or videorecord all their field sessions, so one might suppose that it would be unnecessary to transcribe by hand all the words and sentences provided by the elders. But even aside from the possibility of technological failure, transcription permits you to make notes on variations and on elders' reactions to issues under discussion; and in the Salish-Pend d'Oreille case, the elders like to see the data written on the whiteboard so that they too can write things down. After I've written an item on the board in a spelling that suits everyone, I have to stop to write it down on paper. This slows down the work and helps make the pace leisurely, so it is beneficial overall.

One could also work briefly with several different consultants during the day, thus avoiding any chance of tiring any of them too much, and for some purposes this is a good idea. But because it takes time each day to establish a steady rhythm of the work, switching consultants during the day will inevitably result in loss of time. In addition, and more importantly, there is a major advantage to working with a group (aside from the fact that if one consultant dozes off after lunch,

there are others available to provide data): with a group, you hear table-talk in the language, they jog each other's memories, you get both men's and women's perspectives (assuming your group is mixed), and you also provide the elders with a rare and valuable opportunity to speak their language.

Preparing thoroughly for each field session is vital. You might imagine, if you are starting from scratch with zero knowledge of the target language, that eliciting data will proceed briskly if you just think of things to ask the elders as you go along. Don't count on it. It's fine for your consultants to indulge in free-association thinking of words to offer you – you'll get useful data that way – but you need to be ready to fill all the available time, and to do that you must have a script, a set of questions or other prompts prepared in advance. It now takes me almost two full days to get ready for one day on the reservation, although it took less preparation time thirty years ago, when the things I most needed to hear were paradigms and other beginning-of-documentation items. Being thoroughly prepared basically means having more material than you expect to cover in a single day; that ensures that you won't run out of things to ask about. Some field projects may afford the fieldworker leisure to collect data over many years and with many consultants. But with an endangered language, your time may be quite limited: elderly speakers have linguistic knowledge that might be completely beyond the capacity of younger community members, and no one lives forever.

After you have prepared thoroughly and arrived at your field site, you should also be prepared to have the session canceled without warning. Things happen: the huckleberries may be ripe, so that the elders have gone out into the woods to pick huckleberries; they may have doctor appointments (though working with a group helps to preserve your day's work, because with any luck they won't all have doctor appointments at the same time); they may have visiting children; they may have a new great-grandchild to inspect. And they may get sick. Because so many things can disrupt a scheduled field session, your only choice is to suppress your time anxiety. Don't expect things to happen on a tight schedule; they won't. You can always reschedule a session, as long as you haven't committed yourself to a now-or-never schedule.

The final topic in this section is the question of choosing data-collection techniques (see e.g. Mithun 2001: 34–48 for further discussion of this issue). No fieldwork project can hope to succeed by employing just one technique; but using a variety of ways of gathering data is even more important when the target language is endangered, because of the need to help consultants remember parts of their language that they might not have used for many years. The starting point is always direct elicitation of words, paradigms, simple sentences, and the like ("How do you say 'walk' in your language?"). This enables you to work your way into the language by giving you building blocks for understanding more complex

constructions and concepts. It is by far the fastest way to get lots of words and basic grammatical features at the outset.

But elicitation has pitfalls when you're working with fluent bilinguals, which is almost certainly the situation you'll be dealing with in your research on an endangered language. Sometimes, for instance, a bilingual speaker may give you translations that are as close as possible to the dominant language's structure rather than translations that sound natural in the target language. Some years ago I was eliciting ditransitive sentences like *Johnny stole huckleberries from Mary* from a 98-year-old Salish speaker, and he kept giving me translations like this:

Čoní	naq^{w}'	t	sťša	tl'	Malí
Johnny	steal	OBL	huckleberry	from	Mary

These sentences were startling, because although they were completely grammatical, they were highly marked from a discourse perspective and therefore extremely peculiar as isolated sentences. First, the usual sentential word order in the language is verb-initial; putting the agent NP first indicates special emphasis on that sentence component. Second, the bare-root verb form *naq^{w}'* 'steal' would normally be used only in a limited range of discourse contexts, for instance when there's a change of agent in mid-story (as in the last clause of a sentence like "X did this and X did that and the other thing, and then Y did something") – and obviously no such discourse context exists in an isolated elicited sentence. Normally, a morphologically complex ditransitive verb form would be used as the default in such a sentence.

Finally I asked my consultant, "Joe, aren't these sentences a bit, um, Englishy?" He was surprised. "Yes, of course they are. I thought that's what you wanted – you were asking in English!" "Ah", I said. "So how would you say these things if you were just talking normally in your language?" "Well then," he said, "it'd be like this":

Naq^{w}'-m-ɬ-t-s			Malí	t	Čoní
steal-DERIVED.TRANS-2ndOBJ-TRANS-3.AGENT			Mary	OBL	Johnny
t	sťša				
OBL	huckleberry				

This is an extreme example, but the danger of assimilation of the target language to the language of elicitation is always present with a bilingual consultant.

Elicitation is limited in other ways as well. Some grammatical features may be so rare in a gravely endangered language that speakers can't, or can't easily, produce them on demand. Early in my research on Salish-Pend d'Oreille, for example, I found it almost impossible to get translations of sentences with first-person plural agents acting on second-person plural patients and vice versa – perhaps because it had been a long time since the language was used regularly with large enough groups

of speakers and addressees. So I used a different technique: instead of eliciting isolated sentences, I made up a story that incorporated the desired constructions and elicited that – still sentence by sentence, but with built-in context. With this technique I was able to elicit sentences like *Nem še l qe Pyél qe olqʷšíɬmt qskʷ'l'sncú.* 'we'll help you cook' (lit. 'FUT DEM LOC 1pl Pete 1pl help.2ndOBJ.TRANS.2plOBJ. BACKGRNDAGNT cook') and *Ye pen' tá! Tá qe qsolqʷšíɬlt!* 'Oh no you won't help us!' (lit. 'DEM but NEG NEG 1pl IRREALIS.help.2ndOBJ.TRANS.1plOBJ.BACKGRNDAGNT').

Moving further away from sentence-by-sentence elicitation, you can collect quasi-elicited texts with cultural content, for instance (on the Flathead reservation) 'How to prepare a deer hide' or 'How to bake camas'. You can start the ball rolling with a question like 'What's the first step in preparing a deer hide?', and then continue with 'And what do you do next?' until your consultant has described the entire process. With an endangered language, this technique may be much easier to use than the gold standard of data gathering: naturally-occurring texts (stories, descriptions of traditional activities, conversations, and other genres). Speakers who are no longer used to speaking their language often find it difficult to produce connected texts. The conversations I've heard at my sessions with Salish and Pend d'Oreille elders are mostly brief snippets, not extended talk. This may be partly due to my presence; but the elders have told me that they rarely get together to speak their language under other circumstances nowadays, so that they have few if any other opportunities to speak it.

To collect stories, you need a story-teller. The last Montana Salish elder I worked with who would tell stories died in 2001. There are tape-recordings from 1975 of accomplished story-tellers, but the tape quality is so poor that I need expert assistance to transcribe them, and my current consultants have little patience for that kind of work.

4. Individual variation

This topic requires little space, but it earns a section of its own because of its importance. Every documentation project, whether the target language is endangered or not, involves an enormous amount of variation. The sources of all this variation are diverse: dialect differences, individual differences, a single speaker's different speech and judgments from one moment to the next and from one session to the next, speech errors, register differences (e.g. formal vs. informal speech), and so on. It is sometimes impossible to determine, with an infrequent construction or phonological feature, whether the variation is patterned or random. Ignoring variation by sweeping data under the rug is not a good idea, but writing a description of a language without abstracting away from at least some of the variation is

not feasible. This is a dilemma that all fieldworkers face, and at times none of the available choices are appealing.

If you are working with the last few fluent speakers of an endangered language, dealing with variation is an even more difficult situation. Distinguishing "correct" from "incorrect" data may be impossible, and interspeaker variation may not correlate with any dialect divisions; it is just as likely that the variation arises from individual differences (see Dorian 2010 for an excellent extended analysis of individual variation in a gravely endangered language). It is therefore wise to keep all variants and look for patterns in the chaos. Because the Salish and the Pend d'Oreilles are separate tribes, I initially expected to find dialect differences among speakers. I didn't find any, beyond a few words for objects.

This raises the issue of anonymity. IRBs tend to be very reluctant to permit identification of individual (human) "subjects"; but in the last stages of a language's life, knowing who provided which pieces of data may be vital for an understanding of the language's lexicon and structure. The solution is to get everyone's permission – the IRB's, the speakers', and the speech community's – to identify utterances by the initials of the utterer. This will make it possible for you and future scholars to determine, for example, that certain conflicting grammaticality judgments result from interspeaker variation, while other differences must be attributed to intraspeaker variation (though even these may be patterned, of course).

5. An extended example: Dictionary-making

A vital part of a primary documentation project is the preparation of a dictionary of the language. Particular difficulties arise with an endangered language because of the need for speakers to remind themselves of words they haven't heard for many years – in some cases, as the Salish and Pend d'Oreille elders I work with tell me, words that they themselves never used, but that they used to hear from their parents or grandparents. Some excellent techniques for expanding a dictionary file rapidly, e.g. having groups of consultants brainstorm using lists of semantic domains (Moe 2001), cannot be used with a language so endangered that its remaining speakers are few, elderly, and no longer accustomed to speaking their language regularly.

How many words are "enough" for a dictionary of an endangered (or other) language? There can of course be no definitive answer to this question, but here are two estimates. Michael Krauss has said that 6,000 words are sufficient for a moribund language – that is, a language that is no longer being learned as a first language by children – and 14,000 words for a non-moribund language (p.c. 2001); Terrence Kaufman suggests 4,000–6,000 words as a minimum dictionary size for

a documentation project, and that 10,000 is a good place to stop (2001:275). No matter how many words you collect, though, a truly complete dictionary is not a realistic goal.

After the very first stages of the fieldwork, when you will elicit short(ish) words from a list of basic vocabulary (typically a Swadesh list), you will be able to collect words from both elicited and (if you have any) naturally-occurring texts. But you will also need to use techniques that are specifically designed to produce additional words for your dictionary, including words for documenting culturally relevant terms and concepts. (Here's a strategy you will not use: you won't start with a monolingual dictionary of English, which begins with words like *a, aardvark, aardwolf, ….*)

First, don't stick rigidly to your 200-word Swadesh list. Some of the words on it, like *at*, won't have straightforward translations in your target language, and you don't want to risk making your consultants uncomfortable by demanding a word they don't have or don't know. Other English words might have more than one translation: you will of course want to get all the words your consultants can think of. In Salish-Pend d'Oreille, for instance, asking for a translation of *brother* will produce two words, *sínceʔ* 'younger brother' and *qéws* 'older brother'.

Many English words that aren't on the Swadesh list will also make your consultants think of different words, and they'll be useful additions to your wordlist even if they aren't exact synonyms. An example: Salish-Pend d'Oreille has several words for 'horse', among them *snčłc'á*(*ʔsqa*) (lit., roughly, 'domestic elk') and *x̣ʌ̓'cín* (lit., roughly, 'bite (grass, etc.)'), but also *kʷlaqín* 'strawberry roan', *čpí* 'palomino horse' *i kʷíl* 'bay horse', and so forth.

Organizing lexical elicitation by semantic fields is also useful, e.g. body parts, kin terms, animal and plant names (pictures help with these), ways of walking, color terms (using a color chart), traditional clothing items, and parts of a tipi, in relevant cultures. But in an endangered language, some semantic domains are already likely to be partly lost, no matter how hard your consultants try to remember them; they may never have been part of your consultants' experience.

Direct elicitation can get you to culturally and linguistically interesting items. Salish-Pend d'Oreille, for example, has several different sound-symbolic formations, and the most interesting one involves triple reduplication of the second root consonant, as in *i p'átttt'* 'sound of a cow-pie plopping' (root: *p'át'*), *i kʷíčččč'* 'the creaking sound a tree makes when it's starting to fall' (root: *kʷíč'*), *i cíkʷkʷkʷkʷ* 'little shiny things sparkling, like stars winking or sequins on a jingle-dress' (root: *cíkʷ*).

When you're eliciting words, open-ended follow-up questions are a good idea. For sound-symbolic words, for instance, you can ask whether there are other words for sounds or sights that you haven't yet asked for; and with luck the elders

will remember some. That's how I collected *i méllll* 'a bunch of things going in and out of vision, like when you see running horses through a picket fence'.

As noted above, you will also need as much textual material as your consultants can provide in the time you have together. There will inevitably be many, many words that you won't get from direct elicitation, because you won't think to ask for them and your consultants won't happen to think of them out of context. This means, of course, that you can't compile an adequate dictionary by focusing solely on eliciting words; you also have to collect texts of as many kinds as possible. But waiting for words to occur in texts (especially if the consultants don't easily provide texts) is not an efficient way of building a dictionary rapidly, so you should also make full use of published sources on the same language, on related languages, and/or on nearby unrelated languages (which will likely share some of the natural-world and cultural features). For my work on Montana Salish, important sources have been dictionaries of Spokane (a dialect of the same unnamed language as Salish-Pend d'Oreille; Carlson & Flett 1989) and Colville-Okanagan (which, like Salish-Pend d'Oreille, is a member of the Southern Interior Salishan subbranch of the Salishan language family; Mattina 1987).

But my most important published source by far is a 19th-century Jesuit dictionary of Montana Salish, *Dictionary of the Kalispel or Flat-head Indian Language*, compiled by Gregory Mengarini and other Jesuit missionaries. (Kalispel, spoken in eastern Washington state, is another dialect of the same language as Montana Salish; Flathead is the name given to the Salish tribe of Montana by whites, for mysterious reasons.) This monumental work comprises two volumes and a total of 1,000 pages: 644 in the 'Kalispel-English' volume, 456 in the 'English-Kalispel' volume. The former contains over seven hundred main entries, many of them with dozens of sub-entries.

The Jesuits' dictionary is hard to use, in part because of underdifferentiated orthography – the letter sequence *ko*, for instance, is used to represent the phonemes /k^{w}'/, /q^{w}/, and /q^{w}'/. Some forms in the dictionary are now very rare or are not recognized at all by the current elders; and there are likely to be some outright errors as well, although the overall quality of the dictionary is impressive.

These features make it desirable to re-elicit much of the material in the dictionary, and I've found that this is an excellent way of generating more words. Similar procedures (minus the orthographic difficulties, with luck) can help in dictionary-making with other endangered languages too. Here are a few examples of ways in which re-elicitation can add words to your dictionary. I might ask, "The Fathers' dictionary has a word *chin-chem-echst* 'I join my hands together to receive something'. Do you have that word? No? Then how would you say that?" The elders' response: "*Čn t'íχčstm*, like, you open your hand to receive something." Or I ask, "The Fathers have *es-chs-chisti* 'he's exploring'. Do you have that?"

Them: "Yes, *esč'šč'ísti*, but it means 'he goes into enemy country to scout'." Whatever source you're using, you encourage the elders to free-associate. Me: "The Fathers have *es-chet-us* 'eyes could be screened off'." Them: "No, but there's *nsčetús* 'the opponent, like in a stick-game'." Or this: Me: "The Fathers have *n-pe-us* 'narrow hole, as of a needle'." Them: "No, we say *esnp'eʔús* 'eye of a needle'; and then there's *esp'ʔúps* 'flat butt'." (And then they giggle.)

In sum, creative use of published sources helps you expand your dictionary files quickly, while enlarging your stock of knowledge of the culture that the language expresses.

6. Conclusion

Documenting an endangered language can preserve linguistic and cultural knowledge that would otherwise be lost forever. This fact makes such a project worthwhile – all the hard work, all the incidental frustrations that are an inevitable part of fieldwork. All fieldworkers must keep in mind the fact that there are two audiences for their research, the speech community and world of academic scholarhship. I won't try to argue for the greater importance of one of these audiences over the other, but it is imperative – as everyone (I believe) now recognizes – for the fieldworker to satisfy the speech community as well as academia. Your research results must be accessible to the community, not just to other linguists; and what happens to the data you gather and the analyses you produce is ultimately their decision, not yours.

References

Bowern, Claire. 2008. *Linguistic Fieldwork: A Practical Guide.* Houndmills: Palgrave Macmillan.

Carlson, Barry F. & Flett, Pauline. 1989. *Spokane Dictionary* [Occasional Papers in Linguistics 6]. Missoula MI: University of Montana.

Chelliah, Shobhanna L. & de Reuse, Willem J. 2010. *Handbook of Descriptive Linguistic Fieldwork.* Dordrecht: Springer.

Crowley, Terry. 2007. *Field Linguistics: A Beginner's Guide.* Oxford: OUP.

Dimmendaal, Gerrit J. 2001. Places and people: Feld sites and informants. In Newman & Ratliff, 55–75.

Dorian, Nancy C. 2010. *Investigating Variation: The Effects of Social Organization and Social Setting.* Oxford: OUP.

Everett, Daniel L. 2001. Monolingual field research. In Newman & Ratliff, 166–188.

Kaufman, Terrence. 2001. Two highly effective models for large-scale documentation of endangered languages. In *Lectures on Endangered Languages:2 – from Kyoto Conference 2000 Endangered Languages of the Pacific Rim,* 269–284. Kyoto: ELPR C002.

Margetts, Anna & Margetts, Andrew. 2012. Audio and video recording techniques for linguistic research. In Thieberger (ed.), 13–53.

Mattina, Anthony, compilor. 1987. Colville-Okanagan Dictionary. In *Occasional Papers in Linguistics*, 5. Missoula: University of Montana.

Mengarini, Gregory S.J., Giorda, Joseph, van Gorp, Leopold; Bandini, Joseph & Guidi Joseph. 1877–79. *A Dictionary of the Kalispel or Flat-head Indian Language.* St. Ignatius MT: St. Ignatius Print. (This work is usually referred to as Giorda 1877–79.)

Mithun, Marianne. 2001. Who shapes the record: The speaker and the linguist. In Newman & Ratliff, 34–54.

Moe, Ron. 2001. Lexicography and mass production. SIL: *Notes on Linguistics* 4: 150–156.

Newman, Paul. 2012. Copyright and other legal concerns. In Thieberger (ed.), 430–456.

Newman, Paul & Ratliff, Martha (eds). 2001. *Linguistic Fieldwork.* Cambridge: CUP.

Rice, Keren. 2012. Ethical issues in linguistic fieldwork. In Thieberger (ed.), 407–429.

Sakel, Jeanette & Everett, Daniel L. 2012. *Linguistic Fieldwork: A Student Guide.* Cambridge: CUP.

Samarin, William J. 1967. *Field Linguistics: A Guide to Linguistic Fieldwork.* New York NY: Holt, Rinehart and Winston.

Thieberger, Nicholas (ed.). 2012. *The Oxford Handbook of Linguistic Fieldwork.* Oxford: OUP.

Thieberger, Nicholas & Berez, Andrea L. 2012. Linguistic data management. In Thieberger (ed.), 90–118.

Thomason, Sarah G. 2001. *Language Contact: An Introduction.* Edinburgh & Washington DC: Edinburgh University Press & Georgetown University Press.

Vaux, Bert, Cooper, Justin & Tucker, Emily. 2007. *Linguistic Field Methods.* Eugene OR: Wipf & Stock.

PART II

Case studies in documentation and revitalization of endangered languages and languages in contact

Converb and aspect-marking polysemy in Nar*

Kristine A. Hildebrandt
SIU Edwardsville

This analysis responds to Michael Noonan's call to embrace the messiness and complexities of grammar found in natural language use, continuing the tradition of undertaking rich, deep investigations of a critically endangered, under-documented language (Nar, Tibeto-Burman, Nepal). It is an examination of the polysemy between a set of non-finite and finite markers in Nar. This paper revises Noonan's labeling to better reflect their distribution in varied contexts. Non-finite *-ce* is analyzed as a perfective converb and *-te* is an imperfective converb, as demonstrated via syntactic and semantic properties. In final position, *-ce* is a gnomic perfective aspect marker and final *-te* is a general imperfective aspect marker. These labels more accurately reflect their situational and temporal semantics.

1. Introduction

This is one of the few papers in this volume with a focus on grammatical analysis, in this case an analysis of converb and aspect-marking polysemy in Nar, one of the languages to which Michael Noonan dedicated his professional life. This paper finds its place in this volume because, as noted by Genetti et al. in their editorial introduction to a memorial issue of *Himalayan Linguistics* honoring the life and works of Noonan (and David Watters, also late), he recognized and responded to the so-called "second-shelf status" where endangered language documentation efforts and publications have traditionally languished in (2011: ix).

* This research was sponsored by a Hans Rausing Endangered Languages Documentation grant (SG-0025). I am grateful to Oliver Bond and Carol Genetti for help with this analysis, to Sally Noonan for sharing Mickey's notes, and to the Nar and Phu communities for teaching me about their language. All errors are my own.

Noonan tirelessly advocated for such documentation outputs to receive more mainstream attention within linguistic theory in general, noting that:

> "…We are engaged in a race against time [and against war, poverty, and the good and bad effects of globalization] to preserve as much of the world's linguistic heritage as possible. In running this race, we face, among others, the following three major obstacles: the standard of grammar writing is not uniformly high, there is not a lot of funding to support grammar writing, and the profession does not sufficiently value or support the writing of grammars." (2006: 352)

Nar-Phu very much fits the description of languages referred to in Noonan's quote: it is critically endangered with under 600 speakers combined between the two varieties, and other than the work begun by him before his passing, there has been virtually no other research published on this language until this time. It is undoubtedly a language undergoing rapid shift, (Noonan 2005), as there is a sharp divide between older people, who are fluent, everyday users of Nar, and those younger Nar who have only passive knowledge of the language, or else are exclusive users of Nepali (Indo-European, the dominant and official language of Nepal). It is appropriate therefore, that this volume continues the tradition of representing the documentation of Nar, a language and a community of speakers to which Noonan was very close.

Nar (and Phu) are mutually intelligible variants of the Tamangic sub-grouping of Tibeto-Burman, and are spoken in villages of the same names in Nepal (Ethnologue: NPA, Endoym: *t͡ʃʰypruŋ*). Despite their close similarities, there is enough cross-dialectal variation to warrant separate treatment, (cf. Mazaudon 1997 for a historical phonological perspective) and so this account focuses on Nar only. Current estimated speaker numbers of Nar are at fewer than 400, and Phu has perhaps 200 active speakers. Observations of outward emigration from Nar and Phu villages to Kathmandu or overseas, combined with data from language attitude and usage interviews carried out by the author, and information gleaned from autobiographical texts, suggest that Nar is moribund; the vast majority of fluent speakers are above the age of fifty, and there is extreme disruption in transmission of the language to younger adults and children. Additionally, there is very little published on Nar-Phu, excepting Noonan (2003), Mazaudon (1997), and a sketch of negation strategies in Nar discourse (Hildebrandt & Bond 2011).

This analysis is a response to Michael Noonan's call to 'describe everything' (2008), or to embrace the messiness and complexities of grammar found in settings of natural language use. The goal of descriptive grammar-writing for Noonan (as opposed to documentation for normalization and standardization) is to create an accurate and complete description of a language (or a language

variety), reflecting actual usage. This endeavor necessarily involves incorporation of texts and (transcriptions) of audio-visual recordings of different discourse genres. Noonan himself was well aware that with this type of goal co-exists the possibility of encountering inconsistencies and contradictions in the distribution and function of lexical and functional elements. Rather than selecting the 'most correct' forms from possible 'vernacular' variation for representation, there should instead be an emphasis on working through variation and inconsistencies towards a representation of practical, actual language usage.

This account is an elaborated examination of the polysemy observed between non-finite converbal constructions (by Noonan's term: 'subordinators') and finite aspect marking in Nar. Noonan's original analysis, reconstructable mainly through brief comments in his 2003 sketch grammar of Nar and in unpublished notes, presents the finite system and subordinators as in Tables 1 and 2, and examples of these forms in finite and clause-combining structures are presented in (1) through (4).

Table 1. Finite verb marking in Nar, as described by Noonan (2003)[1]

	DIRECT (SPKR-WITNESSED)		INDIRECT (SPKR-INDIRECT KNOWLEDGE)	
	POSITIVE	NEGATIVE	POSITIVE	NEGATIVE
PAST	V-cin[2]	a-V-i	V-cin mu	a-V-i mu
AORIST	V-ce	a-V-i	V-ce mu	a-V-i mu
PRESENT	V (mu)	a-V mu	V-te[3] momu	a-V-te momu *or* V-te harmu
DURATIVE	V-te mu	a-V-i	V-te momu	a-V-i momu

1. Abbreviations: 3SG third person singular, ASP aspect, COP copula, DEF definite, ERG ergative, GEN genitive, IMPER imperative, INDIR indirect, LOC locative, NEG negative, NOM nominalizer, POT potential, PST past; Nar has four lexical tones: tone /1/ words have a high, falling pitch and are marked with a circumflex (â), tone /2/ words are high-level and are unmarked, tone /3/ words have a breathy phonation and are marked with a voiced glottal fricative before the vowel (ɦa), tone /4/ words have a low-falling pitch with accompanying breathy phonation and are marked with a voiced glottal fricative before a vowel with a circumflex (ɦâ).

2. /c/ is a post-alveolar affricate [t͡ʃ] and the suffix *-cin* is variably [t͡ʃɨ].

3. *-te* alternates with *-pɛ* in present and durative direct/indirect. Noonan notes that *-pɛ* is non-respect, while *-te* indicates higher respect; also, aorist *V-ce kɛ̂* provides a 'future' time reading, while *kɛ̂* alone is emphatic.

Table 2. Non-finite verb marking in Nar, as described by Noonan (2003)

FORM(S)	FUNCTION
-ne	Adverbial, manner, desideratives & abilitatives, carries a 'potential' reading
-ce[4]	Sequential converb, temporal, sequential sense; occasionally, sequential temporal relationships are zero-marked
-te	Determinant nominalizer, with a more completive, more direct reading (to speaker senses)
-pɛ *or* -re	Indeterminant nominalizer, with a more progressive or more ongoing, intermediate sense (to speaker)
-pɛ-re, -te-re	Purposive, with a subject ergative/non-ergative distinction
-pɛ *or* -pi	Relativizer
-re	Conditional, noted by Noonan as rare in occurrence
-re-me, -te-me	Conditionals, where the 2nd form carries a completive sense

(1) Aorist Aspect *-ce*[5]
*ɛ̂le=ce saikul=ce pɦor-**ce** ni-ce.*
boy=DEF cycle=DEF take-**CONV** go-ASP
'The boy, taking his cycle, goes.'

(2) Durative Aspect *-te* (Noonan, personal notes)
*pjuŋ=ko kɦrî râ kɦrî **pakhæ-te** mo mû.*
boy=DEF one goat one **bring-ASP** COP INDIR
'One boy brings one goat (over).'

(3) Sequential Converb *-ce* (Noonan notes)
*ɦota=ce ʧû-**ce** pɦi ʧû mo mû.*
3SG=ERG stay-**CONV** speak stay COP INDIR
'Having sat down, he sits speaking.'

(4) Determinant Nominalizer *-te* (Noonan notes)
*ŋâ=ce lakpe=re ɦlike pɦrî-**te** mraŋ-cin.*
1SG=ERG lhakpa=LOC letter write-**DN** see-PST
'I saw Lhakpa write the letter.'

The forms in the gray-shaded cells in Tables 1 and 2 are the focus of this account. Noonan himself admitted that he was dissatisfied with the 'aorist' analysis, and he suggested that the finite sense of *-ce* is likely derived from the sequential converb *-ce*.

4. The sequential and aorist suffixes have several allomorphs: [t͡ʃe, t͡se, se, ʃe].

5. Unless otherwise indicated, examples are taken from the author's field notes. These texts and the transcriptions/interlinearizations may be accessed through the Endangered Languages Archive (http://elar.soas.ac.uk/deposit/hildebrandt2011narphu).

This paper aims to further detail the connections that Noonan hypothesized with data from Nar discourse with the following goals.

First, I will revise Noonan's labeling to better reflect the distribution of these markers in finite and non-finite constructions and contexts. In particular, I propose that in non-finite contexts *-ce* be analyzed as a perfective converb and *-te* be analyzed as an imperfective converb. Their status as converbs will be demonstrated via a variety of syntactic and interpropositional semantic characteristics. Second, I propose that finite *-ce* be analyzed as a gnomic (universal) perfective aspect marker and finite *-te* as a general imperfective aspect marker. These labels more accurately reflect their situational and temporal semantics in both elicited and discourse data.

This analysis then sets the stage for an elaboration and updating of Noonan's original proposal, that what he terms the 'aorist' marker is derived from the sequential converb. It will be shown that there is polysemy between converbal *-ce* and perfective *-ce*, and similarly between converbal *-te* and imperfective *-te* in finite clauses. Table 3 provides a preview of the main analysis of this account.

Table 3. The functions of *-ce* and *-te* in non-finite and finite positions

Form	Non-Finite (Converb)	Finite (Aspect)
-ce	Codes the bounded temporal nature of the non-finite verb in relation to finite	A gnomic (universal) perfective[6]
-te	Codes the durative nature of the non-finite verb in relation to finite	Main verb action or event has some duration

In this sense then, Nar has a series of two aspectual-type (non-adverbial) converbs, which share overlapping temporal semantics (and a likely historical origin) with the aspect markers. While this grammaticalization path in particular is not frequently observed between converbs and finite verb marking in Tibeto-Burman, it is attested in languages from a wide range of genealogical classifications, such as Spanish (Indo-European), Tamil (Dravidian), Turkish (Altaic), Lezgian (Caucasian) (Haspelmath 1995) and Kamas (Samoyed) (Klumpp 2005). In a similar vein, non-finite nominalization strategies are frequently observed in finite (clause and sentence-final) structures (Matisoff 1972; Noonan 1997; Genetti et al. 2009; DeLancey 2011), demonstrating the possible parallel types of grammaticalization cycles across different languages of Tibeto-Burman.

6. Oliver Bond (pc) notes that the *-ce* perfectives in these data are compatible with a 'gnomic perfective' interpretation (universal fact), as gnomic events are not necessarily perceived in terms of their internal temporal structure. In other languages, like Yucatec Maya (Lucy 1994) and Biblical Hebrew (Waltke & O'Connor 1990), the gnomic perfect contributes temporal senses not unlike what is observed in Nar.

The organization of this paper is as follows. Section 2 contains a brief typological profile of Nar. In Section 3, I first provide an account of the two converbal constructions that are the focus of this account, describing their interpropositional semantics and their relevant syntactic properties. In Section 4, I turn to the syncretic forms *-ce* and *-te* as they mark perfective and imperfective aspects, respectively, in finite clauses. In Section 5 I provide a discussion of the likely historical relation between the converbs and aspect markers. In Section 6 I conclude, including a consideration of the role that a grammatical analysis such as this may play in the larger arena of language endangerment, documentation and preservation.

2. Nar-Phu typological profile

Nar (and Phu) is in many ways a typical Tibeto-Burman language with ergative/absolutive and locative/patient case enclitics, definite/indefinite enclitics, postpositions, and SOV clause-level word order. Sentences are typically characterized by a final, (optionally serialized) verb complex, where the finite verb is marked for tense/aspect via suffixation or a combination of suffixation plus copula periphrasis. An example of this is shown in (5).

(5) Serialized Final Verb Complex (Noonan 2003: 346)
nôkju=ce=ce thuŋ nâpraŋ chaŋ=ce ***pi tê lâ-cin.***
dog=DEF=ERG bee insect nest=DEF **go.fast fall do-PST**
'The dog unwittingly knocked down the beehive.'

Some of the examples in this account are from elicitation (largely from Noonan's materials) but most discourse. The discourse examples are from the following sources in Table 4, all recorded and transcribed with assistance from Nar speakers in 2010. In all cases, they corroborate the translations provided in this analysis:

Table 4. Discourse data sources

Text	Genre, Information
MPI stimuli	Free, but controlled responses to photograph and video prompts
The Pear Story	A monologic video stimulus response; a female speaker provides a running descriptive commentary on activities/events in the 'Pear Story' film (Chafe ed. 1980)
Nar Life Story	A male, ca. 60 years, describes his life in the Nar region. Although largely monologic, he does get prompts and reactions from interlocutors
Yaks	A three-participant conversation about the role of yaks in the socio-economic and agricultural history and life of Nar and Phu

In the examples in Sections 3 and 4, I will gloss all instances of converbal *-ce* and 'determinate nominalizer' *-te* as CONV, and all instances of aspectual *-ce* and *-te* as ASP, no matter the semantics conveyed, in order to allow for a focus on their distribution and functions.

3. Converbs in Nar

As is frequently the case in Tibeto-Burman, clause combination in Nar is done via any one of a range of affixes bound to verb bases that are otherwise non-finite. By non-finite I mean that verbs suffixed with these non-finite (non-finite, subordinating) affixes do not otherwise show the full range of morphology (e.g. tense, aspect, evidentials) that finite (and clause-final) verbs do.

The focus of this paper is on the distribution, formal properties and functions of a sub-set of converbal and aspect markers in Nar. By converb, I invoke the definition embraced by Noonan (1999: 401): "…adverbial subordination by means of a set of specialized non-finite forms". The precise definitions and tests for converbs (*vis-a-vis* clause-chaining, nominalization and participials), including formal syntactic criteria of subordination and the non-finite status of the non-finite verb, are the subject of many debates and discussions (cf. Haspelmath 1995; Nedjalkov 1995; Bickel 1998; Genetti 2011, 2005; Coupe 2007 for detailed discussions of the issues). It is beyond the scope of this paper to consider in detail the debates regarding the notion of 'converb', but basic morpho-syntactic evidence demonstrates in Nar that both suffixes, while themselves conveying aspectual distinctions, are bound to an otherwise bare verb-root, which itself does not show any additional finite morphology (although it can be negated directly with the single negative prefix *ma-*). For example, converbal clauses do not show full aspectual marking and they do not show the indirect (evidential) form *mû*. The converb-marked clause is also grammatically dependent on the finite-marked clause for grammatical relations information. A set of examples in (6) and (7) illustrates these differences between converbal-marked verbs and finite verbs.

(6) (Finite clause, with aspect and indirect marked on final verb, Pear Story)[7]
pjuŋ=ko kɦrî râ kɦrî ***pakhæ-te mo mû.***
boy=DEF one goat one **bring-ASP COP INDIR**
'One boy brings one goat (over).'

7. The morpheme *mo* is multifunctional in Nar: it can be a copular verb alone, and it can occur in a periphrastic construction with a main verb to indicate imperfective progressive aspect. It can also double with the particle *mû* to indicate indirectly witnessed actions/events, and in these structures it is variably [mû.mu] or [mô.mu]; This multifunctional distribution is also found in other Tamangic languages.

(7) (Converb clause, with no aspect/indirect marking, Elicitation)
*ɦota=ce ʈû-**ce** pɦi ʈû mo mû.*
3SG=ERG stay-**CONV** speak stay COP INDIR
'Having sat down, he sits speaking.'

In Example (7) the third person singular subject pronoun *ɦota* 'he/she' is governed by the transitive verb *pɦi* 'speak' in the finite (and final) clause because it carries ergative case-marking. Additionally, the converbal-marked verb (*ʈû* 'stay') lacks any additional morphology for tense, aspect or direct/indirect evidentiality.

It is also beyond the scope of this paper to provide a detailed analysis of the full range of converbs (and other non-finite) markers attested in the language, a matter that will have to wait until a complete reference grammar can be produced.[8] What is important to note here is that this analysis puts forward a revision of Noonan's original label 'determinate nominalizer' for the non-finite suffix *-te*, arguing instead for inclusion of this suffix as another contextual converb with an imperfective aspectual reading.

Converbal –ce

I will discuss converbal *-ce* first, demonstrating its functions as a sequential marker, to express entirety, and to express manner. As noted in the introduction, Noonan describes *-ce* as a sequential converb, marking sequential temporal relations between clauses in the sentence. This relation does indeed exist in data from Noonan's personal notes, as shown in Examples (8) and (9).

(8) (Converb, Sequential Relations Between Clauses)
*cuŋ-**ce** ʈû-w!*
catch-**CONV** hold-IMPER
'Having caught (it), hold (it)!'

(9) (Converb, Sequential Relations Between Clauses)
*JM hlekɛ hlô-**ce** kha-cin.*
JM book study-**CONV** come-PST
'JM, having studied, came.'

In both cases, there is a distinct sequencing sense between the actions or events coded across clauses (e.g. in (8) catching happens before holding, and in (9) studying happens before coming).

8. It should be noted that there are some converbs in Nar that can be considered contextual in their interpropositional semantics between the subordinated and main clause (e.g. they convey temporal or aspectual information), and others that can be considered as specialized converbs (conveying presuppositional or adverbial semantics across clauses).

My corpus reveals an abundance of *-ce*-suffixed converbal constructions, including examples where a sequential temporal sense is similarly coded between clauses, as shown in (10) and (11).

(10) (Sequential Relations, Life Story)
*jarcokʰompa khæ-**ce**, lalemʰi tʰoŋre a-caŋ mo.*
yarcegompa come-**CONV** someone animal NEG-take.care COP
'When yarcegompa[9] season comes, there is nobody (around) to care for (our) animals.'

(11) (Sequential Relations, Pear Story)
*châŋ=ce phjaŋ=re ten-**ce**, rumal=ce râŋlo*
basket=DEF up=LOC put-**CONV**, hanky=DEF again
âŋcaŋ=re kun-ce.
neck=LOC fasten-ASP
'Having put the basket back up, again (he) fastens the hanky (to his) neck.'

In (10) and (11) (also same-subject unless explicitly coded), one event happens or takes place before another. However, there are just as many converbal constructions suffixed with *-ce* where there is no sequential reading available. Consider Examples (12) and (13).

(12) (Yaks)
*kjoloŋ car-**ce**., otare kæ khoreŋ tɛ̂-pa-ni.*
round twist-**CONV** like.this EMPH rope.type become-NOM-POT
'(We) spin/twist (the rope fibers) around. Like this, in this way, it will become the rope-wood.'

(13) (Yaks)
pricuŋ pje khoreŋ, ʰjâ to tʰam-pi=je
female yak rope, yak bundle tie-NOM=GEN
*khoreŋ, ocu pi-**ce** so-pe.*
rope, like.this say-**CONV** make-NOM
'The rope that is tied to the bundle, doing like this, (this is how we) make (that type).'

9. Yarcegompa is a much-coveted fungus, with supposed medicinal qualities, that is harvested by Nar and Phu people for profit in late spring/early summer in high mountain pastures near the Nepal-Tibet border.

(14) (Yaks)
ŋĥî=je, kĥjaco casome hen, a-kĥoŋ kĥom-pe
1PL=GEN, rope compare little.bit, NEG-be.strong be.strong-NOM
*tɑ̂-**ce**, ʃen casome cû kĥom kĥom-pe.*
become-**CONV**, other compare this be.strong be.strong-NOM
'Compared to ours, the other rope type (gotten in Kathmandu) isn't so strong – ours (the local type) is stronger in comparison.'

In (12) through (14), a more accurate semantic interpretation is that the relationship between two or more actions or (attributive) states is viewed in its entirety. The converbal construction functions to highlight the connectedness of actions or events (as a kind of a recapitulator), rather than to frame one as temporarily prior to another. This recapitulation function is frequently observed in Manange, a related language (Hildebrandt 2004: 126). At other times, either sequential or manner readings are both available (in other words, the interpropositional semantics are ambiguous), and this is shown in (15).

(15) (Pear Story)
*ɛ̂le=ce saikul=ce pĥor-**ce** ni-ce.*
boy=DEF cycle=DEF take-**CONV** go-ASP
'The boy, taking his cycle, goes.'

In (15), the interpretation could be that the boy first takes his cycle (up from the ground), and then he goes. Another valid interpretation is that the boy goes, taking his cycle with him as he moves along. As the Nar speaker with whom I was transcribing this text explained to me, both meanings are acceptable in this context.

In still other examples, only a manner reading is conveyed between the converbal clause and the main clause. Examples (16) and (17) illustrate this manner reading more clearly.[10]

(16) (Yaks)
*mjɛ̂-re kju-**ce**, mjɛ̂=re kju-**ce**, mjɛ̂=re kju-**ce**,*
fire=LOC put.in-**CONV**, fire=LOC put.in-**CONV**, fire=LOC put.in-**CONV**,

10. Example (16) also illustrates the tendency for converbs to occur in multiplicity across several non-finite clauses, further blurring the distinction between clause chains and converbs in Nar. Additionally, this example illustrates the possibility for clauses in final position (and interpreted as finite) in Nar to occur with the nominalizer, making Nar another language that supports Delancey's (2011) analysis of the development of nominalizers on final clause constructions in several Tibeto-Burman languages.

*pɦalapuli khûr lɛ̂-**ce**, te=e khûrere lɛ̂-**ce**, car-pe.*
quick spin do-**CONV**, there=LOC spin.spin do-**CONV**, twist-NOM
'Putting (these wool threads) quickly in the fire, like this, repeatedly, and by continuously spinning it, we twist/braid (it).'

(17) (MPI Cut-Break 41)
*sakam cici=ce thôŋ-**ce** then-cin.*
box little=DEF open-**CONV** put-PST
'(The woman) opened the little box wide.'

Example (17) is a response to an MPI video stimulus where a woman opens a hinged lid on a box (similar to on a jewelry box). In a single motion, she opens the lid completely back, but the lid does not separate or become placed on the table. In this sense, the converbal clause plus main clause convey a meaning of 'open wide', rather than 'open and then put', a kind of manner reading and not sequential. In fact, the manner semantics of the *-ce* converbal construction in Example (17) can be better appreciated in light of Example (18), where a true sequential relationship is conveyed between actions via a serialized construction.

(18) (MPI Cut Break 42)
*botal=ce, ko **thôŋ then-cin**.*
bottle=DEF, lid **open put-PST**
'(The woman) opens the bottle lid (and puts it down?).'

In Example (18), the woman opens a different type of bottle and then places the (un-hinged) lid onto the table. These separate actions in this example however are coded without the use of the converbal suffix, and instead with a serial verb structure. Noonan noted that sequential relations could occasionally be zero-marked, and this might be such an instance.

Converbal -te

In comparison to converbal *-ce* are instances of non-finite *-te*. Noonan identifies this as one of a set of nominalizers in Nar-Phu, coding non-finite actions that carry a more completive or more direct reading to the speaker. This nominalizer, according to him, is in complementary distribution with non-finite *-pɛ* (or its allomorph *-rɛ*), which codes actions with a more progressive or ongoing, intermediate sense to the speaker. He provides minimal pairs to illustrate this difference, as in (19) and (20), retaining Noonan's original glosses (IN indeterminant nominalizer and DN determinant nominalizer) for now.

(19) (Elicitation, Indeterminant Nominalizer, Noonan 2003: 346)
*ŋâ=ce lakpe=re ɦlike pɦrî-**pɛ** mraŋ-cin.*
1SG=ERG lhakpa=LOC letter write-**IN** see-PST
'I saw Lhakpa writing the letter.'

(20) (Elicitation, Determinant Nominalizer, Noonan 2003: 346)
*ŋâ=ce lakpe=re ɦlike pɦrî-**te** mraŋ-cin.*
1SG=ERG lhakpa=LOC letter write-**DN** see-PST
'I saw Lhakpa write the letter.'

In my corpus, there are ample instances of non-finite clauses marked with *-te*, and in some instances, the action does carry a more completive (or sequential) sense in relation to the sense coded by the verb in the final clause of the larger sentence. This is shown in (21) and (22).

(21) (Pear Story)
*puci=ce khar-**te** nuŋ kɦola=ce, phjaŋ ʧɦwi-ce.*
knee=DEF brush-**CONV** 3SG cloth=DEF, up pull-ASP
'Brushing off his clothing, he pulls (his clothing) up.'

(22) (Cut Break 04)
*mɦi=ce=ce ɦyetaŋ **khæ** **khæ-te**, katen=ce*
person=DEF=ERG anger **COME** **COME-CONV**, fabric=DEF
ruk tha-cin.
completely cut-PST
'The person, becoming angry, totally cuts up the fabric.'

In many other instances though, the relationship between the verbs in the two clauses is more simultaneous or temporally overlapping in nature, as in (23) through (26).

(23) (Pear Story)
*phjaŋ=re ni-**te** ɦlî=ce krê-ce.*
up=LOC go-**CONV** ladder=DEF climb-ASP
'He then climbs up the ladder.'

(24) (Pear Story)
*kɦola=cuke=ri caŋ-**te** chamjaŋ ʧɦwi-te mo mû oce.*
cloth=PL=LOC put-**CONV** continuously pull-ASP ASP INDIR like.this
'Putting them into his clothes (apron pocket), he continuously picks apples.'

(25) (Pear Story)
*klâŋ klâŋ-**te** ni-te.*
play play-**CONV** go-ASP
'Playing, playing (with the paddle), (they) go (off screen).'

(26) (Yaks)
*pimpa ke jite, ta ke jite, ta=ri phruŋ-**te***
wool cover also, suitable cover also, horse=LOC cover-**CONV**

ta-pe, *ɦjâ=ri* *phruŋ-**te*** *ta-pa* *mo* *mû.*
suitable-NOM yak=LOC cover-**CONV** suitable-NOM COP INDIR
'This (soft) wool covering, it is suitable for covering a horse, it can cover a yak.'

Despite an alternating sequential or simultaneous reading, what does seem to persist across Examples (23) through (26) is a durative, ongoing nature of the non-finite verb in relation to the finite verb. For example, the man's anger persists through the cutting of the fabric, and the climbing of the ladder is part of the motion event of going (back up the tree to return to the picking of fruit).

In summary, while some examples do reflect the functions that Noonan was able to observe, many other examples complicate an analysis of *-ce* as a sequential converb and *-te* as a determinant nominalizer (in complementary distribution with the indeterminant nominalizer). I propose an alternative analysis, that while not drastically different from Noonan's, does increase the number of converbal affixes by one and more accurately reflects the interpropositional semantics coded by non-finite verbs in relation to finite verbs. This is shown in Table 5.

Table 5. Perfective and imperfective converb functions

Form	Function
-ce PERFECTIVE CONVERB	The *bounded temporal nature* of the non-finite verb is highlighted in relation to the finite verb
-te IMPERFECTIVE CONVERB	The *durative* (*sequential or simultaneous*) *nature* of the non-finite verb is highlighted in relation to the finite verb

As for the indeterminant nominalizer *-pɛ*, this marker is rare in occurrence in my corpus in comparison to converbal *-te*. The one example that I have is the following (27).

(27) (Pear Story)
ɛ̂le=ce *khæ-**pɛ**,* *mraŋ-ce.*
old.man=DEF come-IN, see-ASP
'The old man sees (them) coming.'

The overall loose translation into English is similar to that in (19) ('I saw Lhakpa writing the letter.'). This could in fact indicate that it is a nominalizer that renders the non-finite clause as a complement to the main verb. In this sense in (27), what the old man sees is the 'coming (of the boys)', and in (19), what the first-person subject sees is 'the writing of the letter by Lhakpa'. At this point, this remains speculation, and the functions and distributions of *-pɛ* would need more careful examination in comparison with the other nominalizers and relativizers that are frequent in Nar elicitations and discourse.

4. Aspect-Marking

As mentioned, *-ce* and *-te* are also found on the final verb, and are noted by Noonan to code aorist and durative readings of activities and events, respectively. Elicited examples from Noonan's notes support this as shown in (28a–b) and (29a–b).

(28) (Aorist *-ce*)

a. *mɦi=ce nôkju se-**ce** (mû).*
man=ERG dog kill-**ASP** (INDIR)
'The man (just now) killed the dog.'

b. *ŋɛ̂=ce nôkju=re pɦâlto phruŋ-**ce**.*
1SG=ERG dog=LOC leg strike-**ASP**
'I kicked the dog (just now).'

(29) (Durative *-te*)

a. *ŋɛ̂ cɦɛ̂ thûŋ-**te** mo (mû).*
1SG tea drink-**ASP** COP (INDIR)
'I am drinking tea (now).'

b. *JM cɦɛ̂ thûŋ-**te** mo (mû).*
JM tea drink-**ASP** COP (INDIR)
'JM is drinking tea (right now).'

Noonan actually had a small discourse corpus of about four transcribed and glossed texts, but curiously, neither the aorist nor the durative is observed. Rather, the majority of instances of these forms in final position come from his notes of elicited speech.

Aspectual -ce

I will focus on my analysis of *-ce* as an aspect marker first. In my own corpus, I have also found instances where *-ce* codes a past event or action, as in (30).

(30) (MPI Stimuli, Cut-Break #21)
*khatsar=ce ŋɦîtum **pi-ce**, karte=ce.*
carrot=DEF two.piece **cut-ASP**, knife=DEF
'The knife has cut the carrot into two pieces.'[11]

However, many other examples from my corpus challenge any kind of past event or action reading. At times, the verb suffixed with *-ce* may be interpreted as a general

11. Commas in the Nar texts (,) indicate intonation-unit re-sets, while periods (.) indicate the end of a syntactic (sentence) unit. In this example, the incident – the cutting of the carrot – has been directly witnessed by the speaker; hence indirect *mû* is absent.

fact, without reference to any particular time or any notion of (in)completion. This is illustrated with Examples (31) through (33).

(31) (Yaks)
*cû khoreŋ=ce, ta-ni lɑ̂-**ce** mo mû.*
this rope.type=DEF, be.nice-ADV do-**ASP** COP INDIR
'We make this type of wool rope with great care.'

(32) (Yaks)
kun=ce=re tepe mɦlaŋ=ce tenne tar=ce ŋɦî=je
middle=DEF=LOC again black=DEF and white=DEF 1PL=GEN
*pate kjo-ce, njaŋ nimtom mɦlaŋ lɑ̂-**ce**.*
mixed.black.white weave-CONV 1PL opinion black do-**ASP** COP
'Again, the black-and-white (type of wool), wrapped in the middle, I believe we do (wrap) the black.'

(33) (Life Story)
*njema amta njema ʈuja to-**ce.** ta mû.*
1PL.GEN income 1PL.GEN animals need-**ASP** become INDIR
'Our income relies upon our animals. It's become like that (for us).'

In at least one elicited example from Noonan's notes, *-ce* codes a future, non-completed, unrealized and potential reading, as in (34).

(34) (Elicitation)
*ŋɑ̂ khæ-pɛ lho phalpe=re khæ-**ce**.*
1SG come-NOM year Kathmandu=LOC come-**ASP**
'Next year, I'll come to Kathmandu.'

Another example from my corpus is from a Nar speaker's rendition of a scene in *The Pear Story*. A group of boys is assisting another boy who has fallen from his bicycle, while another boy observes and plays with a badminton paddle. In Example (35), *-ce* is suffixed to a light verb that codes ongoing or continuous action, indicating the continuation of playing badminton while other actions in the story line are ongoing simultaneously.

(35) (Pear Story)
*kɦrî-pa=ce na badminton kalpɛ klâŋ ʈû-**ce** mo mû.*
one-NOM=DEF DM badminton like play stay-**ASP** COP INDIR
'One of them continues to play with a badminton (lit: 'has continued to play').'

In Example (35), *-ce* also gives a general reading about the event as a whole, and does not explicitly code any past time or any degree of completion of the action.

Aspectual -te

Turning to the *-te* suffix, some of my corpus examples do correspond with a durational or non-punctual reading of the event in similar fashion to Noonan's analysis, as in (36) and (37).

(36) (Pear Story)
*thôŋpɛ-ce mâr ʧhwi-**te** mo mû.*
tree-DEF down pick-**ASP** COP INDIR
'Down from the tree, (he) is picking (apples).'

(37) (Pear Story)
*teta-ce tepe ʃiŋ pjuŋ-ko kʰrî râ kʰrî pakhæ-**te** mo mû.*
from.that-DEF again wood boy-DEF one goat one bring-**ASP** COP INDIR
'From (over) there (the forest) again, one boy is bringing one goat.'

In other cases the suffix corresponds with a present-time or general imperfective reading of the action or event rather than a specific durative or continuous reading, as in (38) and (39).

(38) (Yaks)
*ʧoŋ-cuke-ri oce, mi a-pu-pɛ oce ŋʰûŋ-**te** hjâŋ.*
forest-PL-LOC like.this, fire NEG-light-NOM like.this sleep-**ASP** opinion
'In the woods, (even if we) cannot light a fire, we can just sleep like this (in our yak-wool blankets, warmly).'

(39) (Pear Story)
*kaʃa-ce, ê̂le-ce krê̂-ri chwɛj-ce ni-**te.***
special.kind-DEF boy-DEF head-LOC put.on-CONV go-**ASP**
A special kind (of hat), the boy puts it on his head and goes (off).'

In summary, while *-ce* can and does code past events, it also codes general facts, or statements about general states of affairs, without explicit reference to past time. It also codes actions that are not yet realized or achieved, but that have possible future relevance. Likewise, while *-te* can and does code durative, ongoing actions, it also codes simple present events and actions, or those actions being witnessed just as they unfold, even if the actions themselves are non-continuous.

5. From converb to aspect marker

In Sections 4 and 5 I have refined and demonstrated the overlapping temporal semantics between the non-finite converbal and the finite aspect suffixes in Nar. These shared semantics are outlined in Table 6.

Table 6. Shared semantics of converb and aspect suffixes

Form	Non-Finite (Converb)	Finite (Aspect)
-ce	Codes the bounded temporal nature of the non-finite verb in relation to finite	Main verb action or event is bounded, but not necessarily complete or homogeneous in terms of its internal structure; also, the internal details of the event structure are not relevant, a gnomic (universal) perfective[12]
-te	Codes the durative nature of the non-finite verb in relation to finite	Main verb action or event has some duration, as opposed to being punctual or an achievement

These overlapping semantics are not accidental, and thus these suffix sets are very likely polysemous, whereby the non-finite use has become acceptable as a grammatical marker on a finite verb in Nar. This situation is also not typologically unique, as Noonan noted that a similar path had occurred for the progressive converb (*-gəy*) in related Chantyal, which in finite clauses now conveys a sense of duration or a generally agreed upon and socially arranged event (1999: 408–409). In Chantyal, however, the sequential converb has not grammaticalized into a perfective aspect marker, but has instead come to be used in main verb constructions coding get-passives, benefactives, reciprocals and adverbial manner of the main verb (411).

A similar situation is observed in languages from other families and geographic locations, for example Kamas (Samoyed, extinct), where converbal constructions can be periphrastic (Klumpp 2005: 398). The non-finite verb is suffixed with the general converb marker *-LAʔ* and a series of phonologically independent auxiliaries may optionally convey additional perfective and imperfective aspectual information about the clause. These auxiliary elements have been subject to reanalysis as aspect markers on main verbs, via a process of phonological erosion and semantic reduction.

Haspelmath (1995: 43–44) notes that the converbal strategy is amenable to the development of a progressive (and presumably an imperfective) reanalysis. He cites Bybee and Dahl (1989: 77), who observe that the source of progressives is converbs, which themselves may have been locatives or locative forms of verbal nouns. They may then be reinterpreted with the semantics of 'at or located at an

12. Oliver Bond (pc) notes that the *-ce* perfectives in these data are compatible with a 'gnomic perfective' interpretation (universal fact), as gnomic events are not necessarily perceived in terms of their internal temporal structure. In other languages, like Yucatec Maya (Lucy 1994) and Biblical Hebrew (Waltke & O'Connor 1990), the gnomic perfect contributes temporal senses not unlike what is observed in Nar.

activity,' setting the stage for imperfective aspect marking on the main verb. As is seen with Nar, these progressive forms are frequently in periphrastic constructions with a (locative) copula. Haspelmath also notes that anterior or perfective converbs used in periphrastic constructions may pave the way for the rise of a resultative construction, in turn setting the stage for the development of perfect aspect. In the case of Nar, a reading of sequentiality is possible with the *-ce* converb construction, but it could also be that the bounded event structure of the converbal clauses as a whole is what contributes to its function as a gnomic or universal perfective in main clauses.

6. Concluding comments

In this account I have updated and revised Noonan's original treatment of a subset of syncretic subordinating and finite markers with data from a wide range of discourse genres. In particular, this expanded set of data, from elaborated monologic and multi-participant discourse, have allowed me to update Noonan's earlier labels of 'sequential converb' and 'determinate nominalizer', which were proposed by him primarily based on elicited data, and have allowed me to propose a set of related forms that code both perfective and (imperfective) durative sense in non-finite and finite clauses.

This analysis still leaves open questions about the functions and plausible sources of other non-finite markers in Nar, of which there are several. Like its equally threatened sister languages within the Tamangic sub-grouping, there is still relatively little known about clause-combining strategies in Nar, particularly in discourse contexts. Michael Noonan's work set the stage for the careful and comprehensive documentation of Nar that this analysis continues.

As mentioned in the introduction, Nar is a language undergoing rapid shift (death). The discourse material that has contributed to this updated analysis in Nar was available only from older speakers who have lived their lives in upper Manang. Younger speakers (who have largely left Nar, and frequently the country) are not able to provide this type of data.

There are some consequences that emerge from this loss (in general) and these limits in speaker participation (in particular) that are worth noting in the context of this account. Firstly, as K. David Harrison has noted, linguists have very little time left to document most of the world's linguistic diversity before it vanishes forever, and endangered languages play a central role in this race against time (2007: 206). Mithun (this volume) comments that one of the results of the escalating situation of language endangerment is that languages (or the appropriate types

of materials or genres of language use) simply may not be available when new questions come up. This study has shown that only by tapping into the rich and complex scenarios of grammar-in-interaction in Nar can we more fully understand what Genetti (2011) describes for Dolakha Newar and other Tibeto-Burman languages as the "intricate interweaving of [morpho-syntactic] structures, creating a syntactic fabric of depth, subtlety and complexity" (6). Based on the observation that these richly complex, clause combining structures are still fully available to older speakers, yet virtually absent in the speech of younger generations, we are likely observing, to borrow and slightly modify the words form Nancy Dorian in her research on East Sutherland Gaelic (1978: 608), a language that is dying with its morpho-syntactic boots on. As such, the situation in Nar is not one of convergence, but rather the rapid loss of the system itself in its entirety. Without an account like this, and the discourse data from which it emerges, questions like those addressed in this paper simply will not be answerable.

A second consequence of the loss and limitations in work on Nar grammar in the context of this account has to do with the prospects of Nar (and Phu, and its sister languages) in future scenarios where language preservation and revitalization (or revival) could become an actionable community concern. Currently, the Nar-Phu people do not have a working model for language preservation and revitalization. However, a multi-year project begun in 2012 and co-directed by the author and elders in Nar-Phu and related communities aims to incorporate methods of documentation and preservation advocated by scholars like Rice (2011), Czaycowska-Higgins (2009) and Leonard and Haynes (2010). This is a model of *community-based data* collection and preservation such that locals play a role in the decisions of *what* constitutes valid language data, *who* contribute(s) those data, and *how* the results and products could be shaped to be useful to both scholars and the community in a variety of contexts.[13] This model includes the gathering and archiving of data similar to those used in this account: language use in everyday interactions, including extended monologic and multi-participant discourse, and invoking topics and issues of relevance to the communities themselves. As noted by Mithun (this volume), this approach to documentation is beneficial both to scholars and also to local language activists, as the results promote the utility of the language in multiple contexts and genres and the roles it plays in socio-cultural organization.

13. National Science Foundation Documenting Endangered Languages grant 1149639 "Documenting the Languages of Manang, Nepal for Local and International Impact"

References

Bickel, Balthasar. 1998. Converbs in cross-linguistic perspective. Review article of Haspelmath, Martin & König, Ekkehard (eds). 1995. *Converbs in Crosslinguistic Perspective*. Berlin: Mouton de Gruyter. *Linguistic Typology* 2: 381–397.

Bybee, Joan & Dahl, Östen. 1989. The creation of tense and aspect systems. *Studies in Language* 13: 51–103.

Chafe, Wallace. 1980. *The Pear Stories: Cognitive, Cultural and Linguistic Aspects of Narrative Production*. Norwood NJ: Ablex.

Coupe, Alec R. 2007. *A Grammar of Mongsen Ao*. Berlin: De Gruyter.

Czaykowska-Higgins, Ewa. 2009. Research models, community engagement, and linguistic fieldwork: Reflections on working within Canadian indigenous communities. *Language Documentation and Conservation* 3(1): 15–50.

DeLancey, Scott. 2011. Finite structures from clausal nominalization in Tibeto-Burman. In *Nominalization in Asian Languages: Diachronic and Typological Perspectives* [Typological Studies in Language 96], Foong Ha Yap, Karen Grunow-Hårsta & Janick Wrona (eds), 343–360. Amsterdam: John Benjamins.

Dorian, Nancy C. 1978. The fate of morphological complexity in language death: Evidence from East Sutherland Gaelic. *Language* 54(3): 590–609.

Genetti, Carol. 2005. The participial construction of Dolakha Newar. *Studies in Language* 29(1): 35–87.

Genetti, Carol. 2007. *A Grammar of Dolakha Newar*. Berlin: De Gruyter.

Genetti, Carol. 2011. The tapestry of Dolakha Newar: Chaining, embedding, and the complexity of sentences. *Linguistic Typology* 15: 5–24.

Genetti, Carol, Yadava, Yogendra, Grunow-Harsta, Karen, Hildebrandt, Kristine & Watters, Stephen. 2011. Editors' introduction to special issue in memory of Michael Noonan and David Watters. *Himalayan Linguistics* 10(1): ix–xii.

Genetti, Carol; Coupe, Alec; Bartee, Ellen; Hildebrandt, Kristine A. & Lin, You-Jing. 2009. Syntactic aspects of nominalization in five Tibeto-Burman Languages of the Himalayan area. *Linguistics of the Tibeto-Burman Area* 31(2): 97–143.

Harrison, K. David. 2007. *When Languages Die: The Extinction of the World's Languages and the Erosion of Human Knowledge*. Oxford: OUP.

Haspelmath, Martin. 1995. The converb as a cross-linguistically valid category. In *Converbs in Cross-Linguistic Perspective*, Martin Haspelmath & Ekkehard König (eds), 1–55. Berlin: Mouton de Gruyter.

Hildebrandt, Kristine A. 2004. A grammar and dictionary of the Manange language. In *Tibeto-Burman Languages of Nepal: Manange and Sherpa*, Carol Genetti (ed.), 2–189. Canberra: Pacific Linguistics.

Hildebrandt, Kristine A. & Bond, Oliver. 2011. Negation in Nar discourse genres. In *Proceedings of the 3rd Workshop on Language Documentation and Linguistic Theory,* Peter Austin, Oliver Bond, David Nathan & Lutz Martin (eds). London: SOAS.

Klumpp, Gerson. 2005. Aspect markers grammaticalized from verbs in Kamas. *Acta Linguistica Hungarica* 52(4): 397–409.

Leonard, Wesley Y. & Haynes, Erin. 2010. Making "collaboration" collaborative: An examination of perspectives that frame linguistic field research. *Language Documentation and Conservation* 4: 268–293.

Lucy, John. 1994. The role of semantic value in lexical comparison: Motion and position in Yucatec Maya. *Linguistics* 32(4–5): 623–656.

Matisoff, James A. 1972. Lahu nominalization, relativization, and genitivization. In *Syntax and Sematics* I, J. Kimball (ed.), 237–257. New York NY: Seminar Press.

Nedjalkov, Vladimir P. 1995. Some typological parameters of converbs. In *Converbs in Cross-Linguistic Perspective*, Martin Haspelmath & Ekkehard König (eds), 97–136. Berlin: Mouton de Gruyter.

Mazaudon, Martine. 1997. An outline of the historical phonology of the dialects of Nar Phu (Nepal). *Linguistics of the Tibeto-Burman* Area 19(1): 103–114.

Noonan, Michael. 1997. Versatile nominalizations. In *Essays on Language Function and Language Type. In Honor of T. Givon*, Joan Bybee, John Haiman & Sandra Thompson (eds), 373–394. Amsterdam: John Benjamins.

Noonan, Michael. 1999. Converbal constructions in Chantyal. In *Topics in Nepalese Linguistics*, Yogendra P. Yadava & Warren W. Glover (eds), 401–420. Kathmandu: Royal Nepal Academy.

Noonan, Michael. 2003. Nar Phu. In *The Sino-Tibetan Languages*, Randy La Polla & Graham Thurgood (eds), 336–352. Richmond: Curzon Press.

Noonan, Michael 2005. Language documentation and language endangerment in Nepal. Lecture at the University of Iceland, Dialog of Cultures. ⟨http://archiv.ub.uniheidelberg.de/savifadok/volltexte/2008/201/pdf/Iceland_Talk_Handout.pdf⟩

Noonan, Michael. 2006. Grammar writing for a grammar-reading audience. *Studies in Language* 30(2): 351–365.

Noonan, Michael. 2008. Describe everything! Or: Description is not the same as standardization. Paper presented at the Workshop on Language Description, Uppsala, Sweden. ⟨https://pantherfile.uwm.edu/noonan/www/Papers.html⟩

Rice, Keren 2011. Documentary linguistics and community relations. *Linguistic Documentation and Conservation* 5: 187–207.

Waltke, Bruce K. & O'Connor, Michael P. 1990. *An Introduction to Biblical Hebrew*. Winona Lake IN: Eisenbrauns.

Grammatical relations in Mixe and Chimariko

Differences and similarities

Carmen Jany
California State University, San Bernardino

The growing documentation and analysis of endangered and other less commonly studied languages has revealed many unique grammatical systems which can not be explained using traditional concepts, such as subject and object. This paper compares two such systems in two different languages: (a) a hierarchical system with direct or inverse alignment in Chuxnabán Mixe and (b) a hierarchical system based on agents and patients in Chimariko. Although the two systems are very distinct, they share several properties and demonstrate how grammatical marking depends on the grammatical, semantic, and pragmatic properties of the arguments in a clause. Overall, this paper illustrates how the study of endangered and even extinct languages contributes to theories defining the nature of grammatical relations.

1. Introduction

There are many different reasons to document endangered languages. Gippert and Himmelmann (2006) mention three reasons. The first, language endangerment, is the most obvious. It is imperative to document the great wealth of linguistic and cultural diversity found in the world's languages before they disappear. This has resulted in the development of a new sub-field within linguistics, language documentation, now a rapidly emerging trend at linguistics conferences and in linguistics publications (Austin & Sallabank eds 2011; Chelliah & de Reuse eds 2011; Evans 2010; Gippert, Himmelmann & Mosel eds. 2006; Grenoble & Furbee eds. 2010; Thieberger ed. 2012), re-prioritizing the discipline of linguistics. The second reason they mention, the economy of research resources, refers to creating properly archived and accessible language documentations which can be accessed by current and future researchers. The third reason, accountability, builds on the second one. Analyses of data should be verifiable and accessible to further scrutiny. As a result, current language documentation practices show a great concern for creating a lasting, properly archived, accessible, and multipurpose record

of a language. This is imperative, as the growing documentation of endangered and less commonly studied languages has started to significantly shape linguistic theory, in particular linguistic typology. New research is constantly leading to the discovery of new linguistic phenomena, as well as to the restatement and refinement of analytical categories (Epps & Arkhipov 2009). As Evans (2010) illustrates, it is dangerous "to talk about 'universals' of language on the basis of a narrow sample that ignores the true extent of the world's linguistic diversity" (p. xvi). This paper reinforces this point by illustrating how the study of endangered and even extinct languages can lead to the discovery of new phenomena in grammatical relations and, thus, contributes to linguistic theory. Similarly, Mithun (this volume) illustrates how language documentation and academic scholarship can strenghten and inform one another. In the two languages studied in this paper, grammatical relations represent complex and unique features contributing to the distinctiveness and individuality of these languages. Thus, understanding how they work in each language is essential not only for linguistic scholarship but also for language revitalization and maintenance efforts.

Traditionally, grammatical relations have been explained relying on the notions of subject and object. However, the growing documentation and analysis of endangered and other less commonly studied languages has revealed many unique grammatical systems which can not be explained on the basis of subjects and objects (Dryer 1997). While some linguists (Comrie 1989; Givón 1995) believe that subject and object are prototype categories that differ somewhat across languages, others (Dryer 1997) argue for a functionalist approach whereby grammatical relations are viewed in terms of their functional, cognitive, and semantic impact rendering the notions of subject and object unnecessary. The latter view has been adopted in this paper.

In traditional terms grammatical relations "are defined by the way in which arguments are integrated syntactially into a clause" (Bickel 2010), i.e. functioning as subjects and objects based on the morphosyntacic properties, such as verb agreement and case assignment, relating them to the clause. Morever, subjects exhibit certain capabilities, such as coreferential deletion in coordinate, adverbial, and complement clauses or coreference with reflexive pronouns. Recent research has identified further syntactic properties defining grammatical relations in particular languages and has shown that these properties do not always converge on a single set of grammatical relations in a language, such as in cases whereby two arguments show different case marking but both yield verb agreement. Particularly challenging to the notions of subject and object are languages whose grammatical systems are based on referential hierarchies, some of which overtly express event direction. Such grammatical systems are frequent in indigenous languages of the Americas. This paper examines and compares two such systems in two unrelated

American indigenous languages (a) a hierarchical system with direct or inverse alignment in Chuxnabán Mixe and (b) a hierarchical system based on agents and patients in Chimariko. It is shown that although the two grammatical systems are very distinct, they share several properties: (1) only one participant is marked or cross-referenced on the predicate (for the most part – see Example 19 and 20) (2) speech-act participants are ranked higher on the hierarchy, (3) first persons show the most formal distinction, and (4) there is no case marking on arguments. Moreover, in both languages the referential hierarchy effects on grammatical marking can be explained from a functional perspective in terms of subjectivity, politeness, affectedness, and topicality. As a result, the two systems show a close integration of syntax and semantics at the level of predicate-argument relations. In addition, the uniqueness and complexity of these two systems may pose some challenges in a language revitalization and maintenance process. Developing a full understanding of the functional driving forces behind these two systems eases these challenges and creates a deeper insight into these languages.

The paper is structured as follows: Section 2 examines grammatical relations in Chuxnabán Mixe (2.1) and Chimariko (2.2). Section 3 compares the two systems and points to differences and similarities between them, seeking functional explanations. Section 4 presents a summary and reiterates the impact of endangered languages on linguistic theory.

2. Grammatical relations in Chuxnabán Mixe and Chimariko

This section examines two different yet similar grammatical systems in two unrelated languages that are very distinct from one another in terms of their linguistic vitality. While Chuxnabán Mixe, although somewhat endangered, is still being spoken in Mexico, Chimariko, a California indigenous language, has been extinct for about a century. Both languages show grammatical systems based on referential hierarchies, but each is unique in its formal properties and fine-grained details.

2.1 Chuxnabán Mixe, an endangered Mexican indigenous language

2.1.1 *Language and data*

Chuxnabán Mixe is a previously undocumented Mixe-Zoquean language spoken by nine hundred people in one village, San Juan Bosco Chuxnabán, located in the eastern midlands of the Mexican southern state of Oaxaca. Following the framework established by UNESCO (Brenziger et al. 2003) and the levels of endangerment discussed in Grenoble and Whaley (2006: 16–19), Chuxnabán Mixe can be assessed as *unsafe*, *vulnerable*, or *at risk*. All community members

speak the language, and there is intergenerational language transmission. However, education and literacy development occur exclusively in Spanish, and bilingualism is on the rise. Speakers, if literate, are only able to read and write in Spanish, although an orthography has been established for the language and is being used in other Mixean villages. Children in San Juan Bosco Chuxnabán attend school up to the 8th grade. With teachers not being Mixe speakers nor community members, all schooling occurs in Spanish. Older children need to continue schooling in another village. Thus, while Chuxnabán Mixe continues to be learnt as a first language by children, the proficiency of these speakers is limited. Language use is also rapidly declining as many young people migrate to the cities or the United States, where they often cease speaking the language. Moreover, the recent access to satellite TV and internet in the village further contributes to the decrease in Mixe use. Currently, Chuxnabán Mixe is only used within the community, mainly at home and for unofficial daily activities. At official gatherings, such as an informational village reunion, as well as at some religious events, both Chuxnabán Mixe and Spanish are spoken. Communication with members from other Mixean communities generally occurs in Spanish, especially if the variety is very distinct. Hence, there are restrictions in terms of use in language domains. Moreover, virtually no literacy materials existed, nor had an official orthography been adopted prior to the start of my documentation project. At the moment, there are no ongoing language maintenance efforts in the village other than my developing text collection, grammar, and dictionary. However, the village elders recognize the need for extensive data collection and the creation of pedagogical and other materials. While the grammatical system of the language still seems to be fully intact, a Spanish influence is noted in the form of loan words. Interestingly, these loan words tend to be fully integrated into the grammar of the language and can thus occur with Mixean morphology. A detailed study of this phenomenon is beyond the scope of this paper.

The data for this paper stems from personal fieldwork including the collection of over sixty oral narratives ranging from one to fifteen minutes in length and the transcription of eleven of these narratives, as well as the elicitation of specific phrases and verb paradigms. The fieldwork in the village was conducted in 2006, 2008, and 2011. In addition, the data was complemented by weekly sessions with a speaker living in the United States.

In general, the documentation of Mixean languages is not extensive, and fine genetic distinctions still remain unclear. The Mixean territory is composed of 290 communities (Torres Cisneros 1997). Each village speaks a different Mixean variety, many of which are mutually unintelligible. In a number of cases it has yet to be determined whether a particular variety represents a distinct language or dialect. Some linguists divide the Oaxaca Mixean family into three main

languages: Lowland Mixe, Midland Mixe, and Highland Mixe (INEA 1994, 1997a, 1997b, 1997c). Wichmann's classification (1995) further subdivides Highland Mixe into North Highland Mixe and South Highland Mixe. Chuxnabán Mixe has been identified by its speakers as Midland Mixe.

At present, there are a few published grammars and dictionaries of Mixean languages (De la Grasserie 1898; Hoogshagen & Hoogshagen 1997; Romero-Méndez 2008; Ruiz de Bravo Ahuja 1980; Schoenhals & Schoenhals 1982; Van Haitsma & Van Haitsma 1976). While most describe person prefixes and their use, there are only very few detailed studies of grammatical relations in Mixean languages. Typologically, Mixean languages are polysynthetic and head-marking, thus encoding most information in the verb system. Case marking is only found with locatives. Moreover, all Mixean languages have noun incorporation, grammatical relations based on referential hierarchies, and an inverse alignment system. The following section examines the grammatical system in Chuxnabán Mixe.

2.1.2 *Grammatical relations*

Grammatical systems based on referential hierarchies reflect a scale in their grammatical marking whereby speech-act participants are ranked higher than third persons, animate entities higher than inanimates, and known entities higher than unknown entities, thus following the animacy hierarchy proposed by Silverstein (1976). In such grammatical systems, for instance, the hierarchy may determine the choice and/or order of person indices on the predicate. Some languages with hierarchical systems also specifically indicate event direction in transitive and ditransitive clauses. Languages that overtly express event direction indicate via morphological markers on transitive predicates whether the agent or the patient in an event is higher ranked. The action goes in the expected direction ('direct') if the agent is higher ranked or against it ('inverse') if the patient is higher ranked. Such grammatical structures are called direct/inverse systems. Since more documentation and grammars of endangered languages have become available, it is obvious that hierarchical and inverse grammatical systems are quite common in the languages of the Americas (Richards & Malchukov eds. 2012; Zavala 2007; Zúñiga 2006, 2008). One such system is found in Chuxnabán Mixe.

2.1.2.1 *Hierarchical system*. In Chuxnabán Mixe, only the most prominent participant in an event is marked or cross-referenced on the predicate. The hierarchy is determined by three factors: (a) grammatical person: 1st > 2nd > 3rd, (b) animacy: animate > inanimate, and for third persons (c) topicality: human/topical > human/non-topical. If all participants in an event are equal on the hierarchy, the actor is cross-referenced rather than the undergoer. The first person, if present in a clause, is always overtly marked. The hierarchy is illustrated in

Examples (1)–(4). The Chuxnabán Mixe person prefixes are summarized in Table 1.

(1) Most prominent participant marked: 1st person prominence

The Opossum and the Jaguar
cham et ntsutsëch ooyës xchu'utsëch jëtu'unëk inääny
cham et ***n****-tsuts-ëch* *ooyës* ***x****-chu'uts-ëch* *jëtu'unëk* *inääny*
now and **1.A**-devour-TAM good **1.O**-devour-TAM like.that he.said
'And now I will eat you, ok, eat me, he said.'

Example (1) shows that first person, if present, is always marked, regardless of the other participants involved in the event. In *ntsutsëch* 'I will eat you' the first person actor *n-* is indexed on the predicate, while in *xchu'utsëch* 'eat me' it is the first person undergoer *x-*.

(2) Most prominent participant marked: 1 > 2 > 3

a. 2 > 1/3 > 1
x*-yuujx-p*
1.O-wake.up-ASP
'You wake me up.'
'S/he wakes me up.'

b. 3 > 2
m*-yuujx-ë-p*
2.O-wake.up-INV-ASP
'S/he wakes you up.'

c. 2 > 3
m*-yuujx-të-p*
2.A-wake.up-PL-ASP
'You wake them up.'

d. 3 > 3'
y*-aa-tsoyky-py*
3.A-CAUS-cure-ASP
'S/he cures him.'

e. 1 > 2
n*-aa-tsoy-kypy*
1.A-CAUS-cure-ASP
'I cure you.'

Examples (2a)–(2e), mostly stemming from the elicitation of verb paradigms, illustrate the person hierarchy which favors first persons over second over third. Only one participant is overtly marked in transitive and ditransitive clauses. If two third persons occur in a transitive clause, the hierarchy is governed by animacy, humanness, and topicality. This is demonstrated in Examples (3a–b) and (4).

(3) a. Most prominent participant marked: Humanness

The King Kondoy
ja'ay täät tpiijtë tmoojtstë iixtëmë niinyë
ja'ay *täät,* ***t****-piij-të* ***t****-moojts-të* *iixtëmë* *niinyë*
person was **3.A**-find-PL **3.A**-wrap-PL like child
'It was a person, they found him, they wrapped him as (they would wrap) any child.'

b. Most prominent participant marked: Humanness

The King Kondoy
täätä yaa'o'ky yaa'o'këkä'ä lakaartë yaa'o'kjä'ä pëtsëmnë
täät ***y***-*aa-'o'k-y* *lakaartë* ***y***-*aa-'o'k-jä'ä* *pëtsëm-në*
then **3.A**-CAUS-die-ASP lizard **3.A**-CAUS-die-ASP leave-ASP
'Then <u>he</u> killed the lizard, <u>he</u> killed it and left.'

In Example (3a), although the child is more topical in the narrative, it is placed lower on the hierarchy as it is not perceived as a human child. In fact, in the narrative the child is born out of a turkey egg. In (3b), the human participant is also the actor and, thus, cross-referenced on the predicate. In the following example, the topic of the narrative, namely King Kondoy, is always cross-referenced on the predicate.

(4) Most prominent participant marked: Topicality

The King Kondoy
pi'kana'k kita'aktë tiyaxtsoy,
pi'kana'k kita'ak-të ***<u>t</u>***-*iyaxtsoy*
children descend-PL **<u>3.A</u>**-call
'The King called the children to descend,'

pi'kana'k niimääjtsk ets mëëtëka kääky tu'uk kachy
pi'kana'k niimääjtsk ets mëëtëka kääky tu'uk kachy
children the.two and with tortilla one basket
'the two children with a basket of tortillas,'

ets pi'kana'k niimääjstk kiita'aktë estëka kääky tu'uk kachy tkay
ets pi'kana'k niimääjtsk kiita'ak-të etsëka kääky tu'uk
and children the.two descend-PL and tortilla one
kachy t-kay
basket 3.A-eat
'and the two children descended. The King ate one basket of tortillas,'

est pi'kana'k tu'uk tjëëny etsëka pi'kana'k ichëkatu'un xëëmëka
est pi'kana'k tu'uk ***<u>t</u>***-*jëëny etsëka pi'kana'k ichëkatu'un xëëmëka*
and child one **<u>3.A</u>**-devour and children like.this always
'and he devoured a child. The children were always sent'

pi'kana'k yaakääxtë mëëtëka kääky ets pi'kana'k tjëëny tu'uk
pi'kana'k y-aa-kääx-të mëëtëka kääky ets pi'kana'k
children 3.S-PASS-send-PL with tortilla and child
<u>t</u>-*jëëny tu'uk*
<u>3.A</u>-devour one
'with tortillas, and he devoured one.'

Table 1 presents the person affixes in independent and dependent constructions. The distinction between the two constructions is treated in 2.1.2.3.

Table 1. Chuxnabán Mixe person markers

	Independent person prefixes						Dependent person prefixes					
Person	**Intrans**	**Trans A**		**Trans O**			**Intrans**	**Trans A**		**Trans O**		
1 SG	–	*n-*		*x-*			*n-*	*n-*		*x-*		
2 SG	*m-*	*m-*		*m-*			*m-*	*x-*		*m-*		
3 SG	–	*y-*		–			*y-*	*t-*		*y-*		
1 PL EXCL	–	*n-*	*–të*	*x-*	*–të*		*n-*	*n-*	*–të*	*x-*	*–të*	
1 PL INCL	–	*n-*	*-ëm*	*x-*	*-ëm*		*n-*	*n-*	*-ëm*	*x-*	*-ëm*	
2 PL	*m-*	*m-*	*–të*	*m-*	*–të*		*m-*	*x-*	*–të*	*m-*	*–të*	
3 PL	–	*y-*	*–të*	–	*–të*		*y-*	*t-*	*–të*	*y-*	*–të*	

2.1.2.2 *Inversion.* When participants in a clause are such that the actor outranks the undergoer on this hierarchy, there is direct alignment. If the opposite occurs, there is inverse alignment (Dryer 1992, 1994; Gildea 1994; Klaiman 1992; Zavala 2000, 2007; Zúñiga 2006), which is indicated with a special morpheme in Chuxnabán Mixe. Inversion can occur in all scenarios: local (i.e. involving only speech-act participants, such as 1 > 2 or 2 > 1), mixed (i.e. involving speech-act participants and third persons, such as 3 > 2 or 3 > 1), or non-local (i.e. involving only third persons). With two third persons, inverse alignment can be used for pragmatic reasons where the undergoer is more topical than the actor. In such cases, inverse alignment can be marked on nouns, as in Kutenai and Algonquian (Dryer 1992, 1994, 1998). In inverse languages, either both or only one participant can be indexed on the predicate and either both or only inverse direction can be overtly marked on the predicate (Zavala 2000, 2007). Inverse languages are predominantly found in the Americas.

In Chuxnabán Mixe, inverse alignment is indicated with a special verbal suffix *-ë*, as shown in Examples (5) and (6). There is no inverse marking on nouns. Moreover, Chuxnabán Mixe does not mark direct alignment, only inverse alignment. In (5) and (6) inverse alignment is used for pragmatic reasons to focus on the undergoer, the children.

(5) Inverse marker *-ë* in 3' > 3

The King Kondoy
niimääjtsk tää ka'ap yëka'atsu'tsnëtë
nii.määjtsk tää ka'ap y-ëka'atsu'tsn-ë-të
DEM.two then no 3.O-chew-INV-PL
'Then he didn't eat them (the two).'

(6) Inverse marker *-ë* in 3' > 3

The King Kondoy
pi'kana'k yaa'anëëmëtë
*pi'kana'k y-aa-'anëëm-**ë**-të*
children 3.O-CAUS-say-INV-PL
'The person spoke to the children.'

With first person, there is no inverse marker, as shown in (7). However, first persons exhibit different person prefixes in actor and undergoer function, *n-* and *x-* respectively, as in (8).

(7) No inverse marker with first persons

a. No inverse marker
x-yuujx-p
1.O-wake.up-ASP
'You wake me up'
'S/he wakes me up.'

b. Inverse marker with 2nd person
*m-yuujx-**ë**-p*
2.O-wake.up-INV-ASP
'S/he wakes you up.'

(8) Different person markers for actor and undergoer 1st person

a. First person actor *n-*
nmaajchpy
***n**-maajch-py*
1.A-grab-TAM
'I grab you.'
'I grab him/her.'

b. First person undergoer *x-*
xmyaajtspë
***x**-myaajts-pë*
1.O-grab-TAM
'You grab me.'
'S/he grabs me.'

As shown in Table 1, the forms of person affixes vary according to whether the clause is intransitive or transitive and whether the predicate occurs in an independent or dependent construction, as is explained in the next section.

2.1.2.3 *Dependent and independent construction.* As in other Mixean languages, in Chuxnabán Mixe all predicates are treated as either independent or dependent, each with its own set of inflectional person markers. Dependency is triggered if a non-core constituent, such as a temporal or locative adverb or a negative participle, precedes the predicate. This is unrelated to subordination and occurs in both main and subordinate clauses. The distinction between dependent and independent construction is shown in Example (9) and (10).

(9) Independent versus dependent construction

a. Independent
maatsyüüjchpy
***m**-aa-tsyüüjch-py*
2.A-CAUS-hurt-ASP
'You hurt him.'

b. Dependent
ka'ap xyaatsyüch
*ka'ap **x**-yaa-tsyüch*
NEG **2.A**-CAUS-hurt
'You don't hurt him.'

(10) Independent versus dependent construction

a. Independent	b. Dependent
The King Kondoy	The Story of a Cacique
yaaʔoʔkjäʔä tää pëtsëmnë	*ko oojts tjach canción*
y-aa-ʻoʔk-jäʔä tää pëtsëm-në	*ko oojts* ***t***-*jac canción*
3.A-CAUS-die-ASP then leave-ASP	when PST **3.A**-learn song
'He killed it and left'.	'When he learned the song'

In Examples (9), the second person actor is indicated with a different prefix in the independent construction and in the dependent one, *m-* and *x-* respectively. Equally, Example (10) illustrates the two different person indices for the third person in the independent and dependent construction, *y-* and *t-* respectively.

2.1.3 *Summary*

In Chuxnabán Mixe only the most prominent participant in an event is cross-referenced on the predicate following a hierarchy based on grammatical person and on animacy, humanness, and topicality if only third persons are involved. In addition, inverse event direction, i.e. if a participant which is lower ranked on the hierarchy acts as the agent in an event, is overtly marked, except if first persons are involved. First persons show different person prefixes in actor and undergoer function and are always overtly marked, if present. Moreover, there are two sets of person markers, one for independent and one for dependent predicate constructions. Table 2 summarizes the grammatical marking in Chuxnabán Mixe.

Table 2. Summary of grammatical marking

Relation	Marking	Inverse marked	Relation	Marking	Inverse marked
1 > 2	1.A	N/A	3 > 1	1.O	no
1 > 3	1.A	N/A	3 > 3'	3.A	N/A
2 > 1	1.O	no	3' > 3	3.O	yes
2 > 3	2	N/A			
3 > 2	2	yes			

A = actor; O = undergoer; N/A = not applicable.

2.2 Chimariko, an extinct Northern California language

2.2.1 *Language and data*

Chimariko is a now extinct California indigenous language. It was once spoken in a few small villages along the Trinity River and parts of the New River and South Fork River in Northern California. The last speaker probably died in the

1940s. However, the language was last actively used in the late nineteen hundreds. Currently, there are no language revitalization efforts for the language.

Published and unpublished materials on the Chimariko language and culture include an early grammatical sketch (Dixon 1910), a recent grammar (Jany 2009), and handwritten notes from data collection sessions by Stephen Powers (1877) and John Peabody Harrington (1921, 1927, 1928), among others (Berman 2001). The main source of data for this work comes from 3500 pages of handwritten field notes collected by John Peabody Harrington in the 1920s and the notes of George Grekoff. Harrington collected elicited sentences, vocabulary, and oral narratives from several speakers. Grekoff examined Harrington's extensive corpus leaving numerous notes and some analyses which have proven useful.

Typologically, Chimariko is a head-marking language, the same as Chuxnabán Mixe. Core arguments are obligatorily indexed on the predicate, and there is noun incorporation. Case-marking occurs only with instruments and companions and some locatives, while other nominal syntactic relations are unmarked. Grammatical relations are based on agents and patients and on a person hierarchy whereby in most cases only one argument is overtly expressed on the predicate. Chimariko is a mainly suffixing synthetic to polysynthetic language. However, personal pronouns are either prefixed or suffixed depending on the verb stem.

2.2.2 *Grammatical relations*

The same as Chuxnabán Mixe, Chimariko reveals a hierarchical grammatical system favoring speech-act participants over third persons. In addition, it shows an agent-patient distinction for first persons. While the hierarchical system is apparent only in transitive clauses, the agent-patient distinction is found in all types of clauses. First persons are obligatorily indexed either as agents or as patients. As discussed below, this points to subjectivity as a motivation for grammar (Scheibman 2002) and to affectedness as a governing factor for the patient category (Mithun 2008). Agent-patient based and hierarchical argument marking has also been reported for a number of other Native American languages in Northern California, on the US Northwest Coast and in Canada, and in the US Southeast (Mithun 1999, 2008). However, each system is unique in its particular features.

2.2.2.1 *Agent-patient system*. Chimariko grammatical relations are based on an agent-patient system for first persons, which is apparent in intransitive, transitive, and ditransitive clauses, as shown in the examples below.

(11) a. Agent-patient system in intransitive clauses

Harrington 020-1118[1]	Harrington 020-1113
noʔot ʔikʼonip	*noʔot tewčhuxanat*
noʔot ʔ-ikʼo-nip	*noʔot tew-**čhu**-xana-t*
1SG **1SG.A**-talk-PST	1SG big-**1SG.P**-FUT-ASP
'I was talking.'	'I am going to be big.'

b. Agent-patient system in transitive clauses

'Woman wanders'
čhušehemdeʔw, ʔawa hida imamda
***čh**-ušehe-m-deʔw ʔawa hida **i**-mam-da*
1SG.P-take-DIR-DER house lot **1SG.A**-see-ASP
'They took me off, I saw lots of houses.'

c. Agent-patient system in ditransitive clauses

'Woman wanders'
ʔiwo hita čhawund amew
*ʔ-iwo hita **čh**-awu-nd amew*
1SG.A-stay lots **1SG.P**-give-PROG food
'I'll stay here, they gave me lots of food.'

Examples (11a–c) illustrate the distinction between first person agents and first person patients which is independent of the transitivity of the clause. In intransitive clauses, this distinction depends on the verb stem. Some stems take agent affixes while others take patient affixes. A few verb stems can take both types of affixes. The affixing pattern, i.e. prefixes versus suffixes, also depends on the verb stem. Second person singular does not distinguish between agent and patient forms, as shown below.

(12) No agent-patient distinction with 2nd person singular

'Fugitives at Burnt Ranch'
makhotaxantinda, kʼotnihu
***m**-akho-ta-xan-tinda kʼot-ni-hu*
2SG-kill-DER-FUT-PROG run.away-IMP.SG-CONT
'He is going to kill you, run away.'

ʔirʔir musunda mamot, kʼotnihu
*ʔirʔir **m**-usu-nda mamot kʼot-ni-hu*
stranger **2SG**-be-ASP 2SG run.away-IMP.SG-CONT
'You are a stranger, run away.'

In Example (12) the second person singular patient is indexed as *m-* in 'he is going to kill you', the same as the second person singular agent in 'you are a stranger'.

1. The first three digits refer to the microfilm reel 020-024, and the last four digits refer to the page on the reel.

The second person plural, however, shows a distinction between agent and patient forms, *qho-/qh-* and *qha-* respectively.

(13) Agent-patient distinction with second person plural

Harrington 020-1126	Harrington 020-1126
qhuk'oʔnan	*qhak'oʔnan*
***qh**-uk'o-ʔna-n*	***qha**-k'o-ʔna-n*
2PL-talk-APPL-ASP	**2PL.P**-talk-APPL-ASP
'You talked to him.'	'He talked to you.'

A system where second person plural but not second person singular affixes differentiate between agent and patient forms appears irregular and confusing. However, second person plural forms are also special in other Northern California languages, such as Karuk, and are used to show respect to elders (Mithun 2008). This distinction is only apparent in transitive clauses with third person actors, as shown in (14).

(14) No agent-patient distinction with 2nd person plural

Harrington 020-1113
mamqhedot tewqhoxanat
mamqhedot *tew-qho-xana-t*
2PL big-**2PL**-FUT-ASP
'You are going to be big.'

Example (14) demonstrates that the agent-patient distinction for second person plural forms occurs only in transitive clauses. Otherwise, a second person plural patient form *qha-* would occur here, given that the verb stem *tew-* 'to be big' requires patient forms, as seen in (11a). Third persons never show a formal agent-patient distinction. Table 3 lists the personal prefixes and suffixes.

Table 3. Chimariko pronominal affixes

	Singular agent	**Plural agent**	**Singular patient**	**Plural patient**
Verbal prefixes				
First person	*y-, ʔ-*	*ya-*	*čh-*	*čha-*
Second person	*m-*	*qh-*	*m-*	*qha-*[1]
Third person	*h-*	*h-*	*h-*	*h-*
Verbal suffixes				
First person	*-ʔ (i)*	*ya-*	*- čh*	*- čha*
Second person	*-m*	*-qh*	*-m*	*-qh*
Third person	*-h/Ø*	*-h/Ø*	*-h/Ø*	*-h/Ø*

[1]Occurs only in transitive sentences with third person actors.

2.2.2.2 *Hierarchical system*. The same as with Chuxnabán Mixe, in Chimariko only one argument is cross-referenced on the predicate following a hierarchy whereby speech-act participants outrank third persons. The hierarchical system is only apparent in transitive and ditransitive clauses, as illustrated below.

(15) Hierarchical system: 1>3 => 1; 3>1 => 1

'Fugitives at Burnt Ranch'
phaʔasitaʔče yekhotinda, čhaxaduʔxakon, wisseeda čhumčaxa
phaʔasitaʔče ***y****-ekho-tinda* ***čha****-xaduʻx-akon wisseeda čhu-m-čaxa*
that.why 1SG.A-kill-PROG 1PL.P-?-FUT downstreamIMP.PL-DIR-CPL
'That's why I killed him, they will kill us, you all move down to B. Noble's place.'

(16) Hierarchical system: 3>3 => 3; 1>3 => 1

'Crawfish'
hiničxeʔkut, ʔičiʔta, puqhela ʔitxaʔmat
h*-iničxeʔku-t* *ʔ-ičiʔta* *puqhela* *ʔ-itxaʔma-t*
3-smell-ASP 1SG.A-catch basket 1SG.A-put-ASP
'They smelled it, I caught them, I put them in a basket.'

(17) Hierarchical system: 3>2 => 2

'Dailey chased by the bull'
moxowetnan, phaʔyit phuncarye
mo*-x-owet-na-n* *phaʔyit* *phuncar-ye*
2SG-NEG-hook-NEG-ASP thus.say woman-POSS
'He didn't hook you, thus said his wife.'

(18) Hierarchical system: 2>3 => 2; 3>2 => 2

a. *mokoxanaʔ*
m*-oko-xana-ʔ*
2SG-tattoo-FUT-Q
'Are you going to tattoo her?'

b. *qhakʼoʔnaʔ*
qh*-akʼo-ʔna-ʔ*
2PL-talk-APPL-Q
'Was he talking to you?'

Example (15) shows the person hierarchy: when a first person acts on a third, as in *yekhotinda* 'I killed him', the first person is marked; the first person is also marked in *čhaxaduʔxakon* 'they will kill us', when a third person acts on a first. Examples (16)–(18) also demonstrate the person hierarchy in mixed scenarios, i.e. in events involving third persons and speech-act participants. Events with only speech-act participants, i.e. local scenarios, show two different patterns. If a first person is the agent, only the first person is indexed. If a second person is the agent, both the agent and the patient are cross-referenced on the predicate. This is the only instance when two person affixes occur on the predicate, as shown below.

(19) Hierarchical system: 2>1 => 2 + 1 undergoer; 2>3 => 2

a. *mexota*
m-e-xota
2SG-1SG.P-look.at
'<u>You</u> look at <u>me</u>.'

b. *mixota*
m-ixota
2SG-look.at
'<u>You</u> look at <u>it</u>.'

c. *mekhoxana'*
m-e-kho-xana-'
2SG-1SG.P-kill-FUT-Q
'Are <u>you</u> going to kill <u>me</u>?'

d. *makhoxana'*
m-akho-xana-'
2SG-kill-FUT-Q
'Are <u>you</u> going to kill <u>him</u>?'

Example (19) shows the difference between events involving first person patients, as in (19a) and (19c), where both participants are overtly marked, and third person unmarked patients, as in (19b) and (19d). It has to be noted, though, that in (19a) and (19c) the first person undergoer affix *e-* is different in shape from the first person patient prefix *čh(a)-* in instances where only one participant is marked. Following this hierarchy, first persons are always overtly indexed on the predicate, regardless of other participants involved. This is also true in imperative constructions, as in (20).

(20) Hierarchical system in imperative constructions:
2SG > 1; 2SG > 3; 2PL > 1; 2PL > 3

a. Harrington 020-1125
nek'o'na
n-e-k'o-'na
IMP.SG-**1SG.P**-talk-APPL
'Talk to me!'

b. Harrington 020-1125
nik'o'na
n-ik'o-'na
IMP.SG-talk-APPL
'Talk to them!'

c. Harrington 020-1126
čhak'o'na
čh-a-k'o-'na
IMP.PL-1PL.P-talk-APPL
'Talk to us!'

d. Harrington 020-1126
čhuk'o'na
čh-uk'o-'na
IMP.PL-talk-APPL
'Talk to him!'

In (20a) and (20c), the imperative prefix is followed by the first person undergoer prefix *e-* or *a-* indicating the fact that the undergoer of the action of the command is a first person, while in (20b) and (20d) where the command is directed towards a third person, the undergoer is unmarked. The difference between the stem-initial vowel in (20b) and (20d) is due to a morphophonemic process.

2.2.3 *Summary*

Chimariko has a very unique grammatical system with an agent-patient distinction and a referential hierarchy favoring speech-act participants over third persons, thus following the widely established animacy hierarchy (Silverstein 1976). First

persons are treated in a special way. Only first persons show an agent-patient distinction in all types of clauses and always surface, either as agents or as patients or as undergoers in events with second person agents. In intransitive clauses the agent-patient distinction depends on the verb stem. Due to lexicalization and semantic change, a clear patient category involving affectedness, involuntary actions, or the lack of control is no longer observable for the verb stems with patient markers, although many describe actions or states where the participant has no or limited control and is affected (see Mithun 1991). Predicates with patient indices include actions, such as as *give a warcry, cry out, yell (animal), blink, grow up, fall, and sneeze* and others, while predicates with patient markers include states, such as *be called, be mad, be old, be pregnant, be exhausted, be angry, be soft, be decayed, be black, be red*, and others. A small number of verb stems can take either agent or patient affixes. In addition to first persons, second person plural forms make a distinction between agents and patients, but only in transitive and ditransitive clauses. The grammatical system in Chimariko is summarized in Table 4. Given that the language is no longer spoken and the data is limited, all the unique particularities of the grammatical system may never be fully understood.

Table 4. Summary of Chimariko grammatical system

Actor > Undergoer	**Affix on predicate**
1 > 1	**1** agent
1 > 2	**1** agent
1 > 3	**1** agent
2 > **1**	2 + **1** undergoer[1]
2 > 2	2
2 > 3	2
3 > **1**	**1** patient
3 > 2SG	2
3 > 2PL	2PL patient
3 > 3	3

[1]The affix for the first person undergoer is different from the first person patient form.

3. Comparison: Similarities and differences between the two systems

The two grammatical systems described above differ in many aspects, each possessing its own unique features. Nevertheless, they also share many properties, which can be explained in functional terms. As noted in the intoduction, both exhibit

the following: (1) in general, only one participant is marked or cross-referenced on the predicate, (2) speech-act participants are ranked higher on the hierarchy, (3) first persons show the most formal distinction, and (4) there is no case marking on arguments. There is one exception to the first property. In Chimariko, both the agent and the patient are cross-referenced on the predicate in instances where a second person acts on a first person. Both, Chimariko and Chuxnabán Mixe, largely follow the widely established animacy hierarchy (Silverstein 1976) whereby speech-act participants are ranked higher than third persons. However, each language also exhibits special rankings. While Chuxnabán Mixe ranks speech-act participants as 1 > 2 and further ranks different types of third persons, such as animate > animal and human/topical > human/non-topical, Chimariko does not. In fact, in Chimariko third person affixes are phonologically least prominent and sometimes even zero, and speech-act participants are ranked in a special way: the second person plural outranks the second person singular. Similar rankings also occur in Karuk, another California indigenous language. Karuk ranks second person plural highest: 2PL > 1 > 2SG > 3 (Mithun 2008). Mithun (2010, 2012) examines a number of California indigenous languages and notes for Pomoan that speakers use second person plural forms for respect, in particular with elders. Examining further languages, such as Karuk, she concludes (2012: 290) that the status given to second persons "could reflect common politeness behavior". The special ranking found in Chimariko could therefore be explained in terms of politeness. The hierarchies for Chuxnabán Mixe and Chimariko are illustrated below.

Chuxnabán Mixe Hierarchy:	1 > 2 > 3 > 3'
Chimariko Hierarchy:	1 > 2PL > 2SG > 3

Whereas the two languages differ in their hierarchies and in some other features, they share many properties when it comes to first person indexing. In both languages, first persons are always marked on the predicate, regardless of their semantic role or of the other participants involved in an event. Moreover, first person affixes distinguish between actor and undergoer or agent and patient in both languages. Systems where first persons are ranked highest, are always marked, or show the most distinctions, can be explained in terms of subjectivity as a motivation for grammar (Scheibman 2002). Scheibman (2002) examines subjective expressions in naturally-occuring English discourse and notes that in actual language use grammatical structures function more to indicate the speaker's point of view, rather than to provide propositional information. Such expressions are also most frequent in discourse. If first person is the most frequently used in discourse, it should be the most prominent and show the most distinctions. Nevertheless, the opposite explanation is also possible. As Du Bois (1985: 362) notes "grammars code best what speakers do most" referring to the idea that grammars provide the most economical coding mechanism for the most frequently used speech

functions in discourse. The economical coding strategy certainly applies to both languages, as only one participant is indexed on the predicate, with one exception in Chimariko where a second person acts on a first. Furthermore, both languages lack case marking for core arguments.

Several more differences between the two systems need to be pointed out. In Chuxnabán Mixe, there are no agent-patient distinctions in intransitive clauses, and single arguments in intransitive clauses are marked with the same verbal prefixes for all verb stems. Chimariko, on the other hand, shows an agent-patient distinction for first persons extending to all types of clauses. Such systems are also found in other American indigenous languages, such as Haida (Mithun 1999) and Karuk (Mithun 2008) and can be explained by having affectedness as a determining factor for the patient category. Prototypical patients are not in control but rather affected by an event. If affectedness is a relevant factor for the patient category, it only makes sense to apply it solely to first persons, as speakers are unable to evaluate affectedness of hearers or third persons. Another major distinction between the two grammatical systems lies in the marking of event direction. Chuxnabán Mixe overtly marks event direction, while Chimariko does not.

Overall, the patterns found in both grammatical systems illustrate how structure is embedded in function. Both differences and similarities between the two grammatical systems can be explained in terms of their functions, such as subjectivity, politeness, affectedness, and topicality. These functions are certainly not unique to the two languages or to endangered languages in general. However, the study of endangered and even extinct languages may demonstrate new structural ways in which these functions can be fulfilled. Furthermore, understanding these functions in each language and how they interact with the grammatical systems can be useful for language revitalization and maintenance efforts. Mithun (this volume: 40) notes that "structure is intimately connected with substance and context", and this structure represents "part of the knowledge of first language speakers". The two grammatical systems studied in this paper are the result of speakers' choices to encode certain functions and, therefore, show crystallizations of language use.

4. Conclusions

While hierarchical systems seem to be widespread in the Americas as more documentation of endangered languages becomes available, each system is unique in its fine-grained details. The comparison of two such grammatical systems has revealed several similarities which can be explained from a functional point of view. First person prominence points to subjectivity as a motivation for grammar. The distinction of first person agents and patients in Chimariko or actors and undergoer in Chuxnabán Mixe indicates affectedness as a controlling factor

for the patient category. Likewise, the differences between the two systems are open to functional explanations. The special treatment of the second person plural in Chimariko could be related to politeness, as shown in other California indigenous languages, and the distinction between different types of third persons in Chuxnabán Mixe is defined by topicality.

The patterns found in Chuxnabán Mixe and Chimariko demonstrate how grammatical marking depends on the grammatical, semantic, and pragmatic properties of the arguments in a clause, organized hierarchically, and, thus, a close integration of syntax and semantics at the level of predicate-argument relations. The described patterns are unique and very distinct from those widely studied and found in European and other languages. As a result, this paper shows how the study of endangered and even extinct languages can contribute to theories defining the nature of grammatical relations. Such theories go beyond the traditional notions of subject and object or the semantico-syntactic macroroles S, A, and O.

Likewise, academic scholarship can inform language revitalization and maintenance efforts. Understanding the unique grammatical systems and the motivations behind them is essential to developing efficient pedagogical materials for these languages and to creating successful language revitalization and maintenance programs. Moreover, for community members an appreciation for the complexity and uniqueness of the language may lead to recognizing the need for language maintenance and to a higher self-esteem.

To conclude, both linguistic scholarship and community-based language revitalization and maintenance efforts can support one another. On one hand, the distinctive and unique grammatical features examined in this paper contribute to shaping current linguistic theory. On the other hand, their understanding helps guide community revitalization and maintenance efforts.

References

Austin, Peter K. & Sallabank, Julia (eds). 2011. *The Cambridge Handbook of Endangered Languages*. Cambridge: CUP.

Berman, Howard. 2001. Chimariko linguistic material. In *The Collected Works of Edward Sapir* XIV*: Northwest California Linguistics*, Victor Golla & Sean O'Neill (eds), 1039–1076. Berlin: Mouton de Gruyter.

Bickel, Balthasar. 2010. Grammatical relations typology. In *The Oxford Handbook of Language Typology*, Jae Jung Song (ed.), 399–444. Oxford: OUP.

Brenzinger, Matthias; Yamamoto, Akira; Dwyer, Arienne; De Graaf, Tjeerd; Grinevald, Colette; Krauss, Michael; Miyaoka, Osahito; Ostler, Nicholas; Sakiyama, Osamu; Villalon, Marai & Zepeda, Ofelia. 2003. *Language Vitality and Endangerment*. Paris: UNESCO Intangible Cultural Unit, Safeguarding Endangered Languages. ⟨http://www.unesco.org/culture/ich/doc/src/00120-en.pdf⟩ (1 July 2010).

Chelliah, Shobhana L. & de Reuse, Willem J. (eds). 2011. *Handbook of Descriptive Linguistic Fieldwork*. Dordrecht: Springer.

Comrie, Bernard. 1989. *Language Universals and Linguistic Typology*. Chicago IL: The University of Chicago Press.

Dixon, Roland B. 1910. The Chimariko Indians and language. *University of California Publications in American Archeology and Ethnology* 5(5): 295–380.

Dryer, Matthew S. 1992. A comparison of the obviation systems of Kutenai and Algonquian. In *Papers of the Twenty-Third Algonquian Conference*, William Cowan (ed.). 119–163. Ottawa: Carleton University.

Dryer, Matthew S. 1994. The discourse function of the Kutenai inverse. In T. Givón (ed.). 65–99.

Dryer, Matthew S. 1997. Are grammatical relations universal? In *Essays on Language Function and Language Type. Dedicated to T. Givón*, Joan Bybee, John Haiman & Sandra A. Thompson (eds). 115–143. Amsterdam: John Benjamins.

Dryer, Matthew S. 1998. Obviation across clause boundaries in Kutenai. In *Studies in Native American Linguistics IX, University of Kansas Working Papers*, John Kytle, Hangyoo Khym & Supath Kookiattikoon (eds), 33–51. Lawrence KS: University of Kansas.

Du Bois, John. 1985. Competing Motivations. In *Iconicity in Syntax* [Typological Studies in Language 6], John Haiman (ed.). 343–365. Amsterdam: John Benjamins.

Epps, Patience & Arkhipov, Alexandre. 2009. *New Challenges in Typology: Transcending the Borders and Refining the Distinctions* [Trends in Linguistics 217]. Berlin: Mouton de Gruyter.

Evans, Nick. 2010. *Dying Words: Endangered Languages and What They Have to Tell Us*. Hoboken NJ: Wiley-Blackwell.

Gildea, Spike. 1994. Semantic and pragmatic inverse: 'Inverse alignment' vs. 'inverse voice' in Carib of Surinam. In T. Givón (ed.), 187–230.

Gippert, Jost, Himmelmann, Nicholas & Mosel, Ulrike (eds). 2006. *Essentials of Language Documentation*. Berlin: Mouton de Gruyter.

Givón, T. 1995. *Functionalism and Grammar*. Amsterdam: John Benjamins.

Givón, T. (ed.). 1994. *Voice and Inversion* [Typological Studies in Language 28]. Amsterdam: John Benjamins.

Grasserie, De La Raoul. 1898[1968]. *Langue Zoque et langue mixe: Grammaire, dictionnaire* [Biliothèque Linguistique Américaine, Tome XXII]. Paris. Reprint: Kraus. Nendeln, Liechtenstein.

Grekoff, George. 1950–1999. Unpublished notes on various topics. Survey of California and other Indian languages. University of California, Berkeley.

Grenoble, Lenore A. & Furbee, N. Louanna. 2010. *Language Documentation: Practice and Values*. Amsterdam: John Benjamins.

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

Harrington, John Peabody. 1921–1928. Fieldnotes. Microfilm reels: Northern California 020-024.

Hoogshagen Noordsy, Searle, & Hoogshagen, Hilda Halloran. 1997. *Diccionario Mixe de Coatlán Oaxaca*, 2nd edn. Hidalgo, Mexico: Summer Institute of Linguistics.

INEA (Instituto Nacional para la Educación del los Adultos). 1994. *Ayuk Ja'jy Y'ayuujk. Libro del Adulto, Población Mixe (Parte Baja), Oaxaca*. ⟨http://cdi.gob.mx/index.php?id_seccion=519⟩ (1 July 2010).

INEA (Instituto Nacional para la Educación del los Adultos). 1997a. *Ayuk Ja'jy Y'ayuujk. Libro del Adulto, Población Mixe (Parte Alta), Oaxaca*. ⟨http://cdi.gob.mx/index.php?id_seccion=519⟩ (1 July, 2010).

INEA (Instituto Nacional para la Educación del los Adultos). 1997b. *Instructivo del Alfabetizador. Población Mixe, Oaxaca*. ⟨http://cdi.gob.mx/index.php?id_seccion=519⟩ (1 July 2010).

INEA (Instituto Nacional para la Educación del los Adultos). 1997c. *Já'ájy m'ayuujk*. Cuaderno de Ejercicios, Población Mixe, Oaxaca.

Jany, Carmen. 2009. *Chimariko Grammar: Areal and Typological Perspective* [University of California Publications in Linguistics 142]. Berkeley CA: University of California Press.

Klaiman, M.H. 1992. Inverse languages. *Lingua* 88: 227–261.

Mithun, Marianne. 2013. What can revitalization teach us about documentation? In *Responses to Language Endangerment. In Honor of Mickery Noonan. New Directions in Language Documentation and Language Revitalization*, Elena Mihas, Bernard Perley, Gabriel Rei-Doval & Kathleen Wheatley (eds), 21–41. Amsterdam: John Benjamins.

Mithun, Marianne. 2012. Core argument patterns and deep genetic relations: Hierarchical systems in Northern California. In *Argument Structure and Grammatical Relations: A Cross-linguistic Typology* [Studies in Language Companion Series 126], Pirkko Suihkonen, Bernard Comrie & Valery Solovyev (eds), 257–294. Amsterdam: John Benjamins.

Mithun, Marianne. 2010. Contact and North American languages. In *The Handbook of Language Contact*, Raymond Hickey (ed). 673–694. Wiley-Blackwell.

Mithun, Marianne. 2008. The emergence of agentive systems. In *The Typology of Semantic Alignment Systems*, Mark Donohue & Søren Wichmann (eds). Oxford: OUP.

Mithun, Marianne. 1999. *The Languages of Native North America*. Cambridge: CUP.

Mithun, Marianne. 1991. Active/agentive case marking and its motivations. *Language* 67(3): 510–564.

Powers, Stephen. 1877. Tribes of California. In *Smithsonian Institution: Contributions to North American Ethnology* 3, 439–613. Washington DC: Smithsonian Institution.

Richards, Marc & Malchukov, Andrej (eds). 2012. *Linguistische Arbeitsberichte,* 86: *Scales*. Leipzig: University of Leipzig.

Romero-Méndez, Rodrigo. 2008. A Reference Grammar of Ayutla Mixe (Tukyo'm Ayuujk). Ph.D. dissertation, University of Buffalo NY.

Ruiz de Bravo Ahuja, Gloria. 1980. *Mixe Tlahuitoltepec, Oaxaca*. Archivo de Lenguas Indígenas de México. (Data by Don D. Lyon).

Scheibman, Joanne. 2002. *Point of View and Grammar: Structural Patterns of Subjectivity in American English Conversation* [Studies in Discourse and Grammar 11].Amsterdam: John Benjamins.

Schoenhals, Alvin & Schoenhals, Louise C. 1982. *Vocabulario Mixe de Totontepec*. Hidalgo, Mexico: Summer Institute of Linguistics.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In *Grammatical Categories in Australian Languages*, Robert M.W. Dixon (ed.). Canberra: Australian Institute of Aboriginal Studies.

Thieberger, Nicholas (ed.). 2012. *The Oxford Handbook of Linguistic Fieldwork* [Oxford Handbooks in Linguistics]. Oxford: OUP.

Torres Cisneros, Gustavo. 1997. *Mixes*. Comisión Nacional para el Desarrollo de los Pueblos Indígenas. ⟨http://www.cdi.gob.mx⟩ (1 July, 2010).

Van Haitsma, Dieterman, Julia & Van Haitsma, Willard. 1976. *A Hierarchical Sketch of Mixe as Spoken in San José El Paraíso,* No. 44. Hidalgo, Mexico: Summer Institute of Linguistics.

Wichmann, Søren. 1995. *The Relationship among Mixe-Zoquean Languages of Mexico.* Salt Lake City UT: University of Utah Press.

Zavala Maldonado, Roberto. 2000. Inversion and Other Topics in the Grammar of Olutec (Mixe). Ph.D. dissertation, University of Oregon.

Zavala Maldonado, Roberto. 2007. Inversion and obviation in Mesoamerica. In *Endangered Languages*, Peter K. Austin & Andrew Simpson (eds). 267–306. Hamburg: Helmut Buske.

Zúñiga, Fernando. 2006. *Deixis and Alignment. Inverse Systems in Indigenous Languages of the Americas* [Typological Studies in Language 70]. Amsterdam: John Benjamins.

Zúñiga, Fernando. 2008. How many hierarchies, really? Evidence from several Algonquian languages. *Linguistische Arbeitsberichte,* 86: *Scales,* 277–294.

List of abbreviations

1	First person	IMP	Imperative
2	Second person	INV	Inverse
3	Third person	NEG	Negative
A	Agent/actor	O	Undergoer
ASP	Aspect	P	Patient
APPL	Applicative	PASS	Passive
CAUS	Causative	PST	Past tense
CPL	Completive	PL	Plural
CONT	Continuative	POSS	Possessive
DEF	Definite	PROG	Progressive
DEM	Demonstrative	S	Intransitive argument
DER	Derivational	SG	Singular
DET	Determiner	TAM	Tense/Aspect/Mood
DIR	Directional	Q	Interrogative
FUT	Future		

Having a *shinshii/shiishii** 'master' around makes you speak Japanese!

Inadvertent contextualization in gathering Ikema data**

Toshihide Nakayama & Tsuyoshi Ono
Tokyo University of Foreign Studies / University of Alberta

Everyone brings with them a particular set of contextualization to interaction. After several years of working on the Ikema dialect of Miyako Ryukyuan, we are finally realizing that our very presence as *shinshii/shiishii* 'masters' is one of the major factors encouraging Ikema people to use Japanese. Because of the way our social identity is perceived among the Ikema people, our work session turns out to be an inherently a Japanese speaking context, which not only makes collection of naturalistic discourse data rather precarious but also can influence even the native speakers' grammatical intuition of Ikema. Our goal in this paper is to explicate the very sensitive nature of the fieldwork setting and its implication for data collection by illustrating how the identity of researchers themselves can inadvertently shape the context and in turn shape the choice and use of language in fieldwork situations.

1. Introduction

Traditionally, when linguists think of issues of data collection in linguistic fieldwork, attention tends to be given more toward what forms and constructions to collect and how to collect them than how speakers behave linguistically (i.e. how

---

* The Ikema form which means 'master' is *shinshii* on Ikema Island and *shiishii* in Nishihara on Miyako Island.

** This research initially involved Shoichi Iwasaki who was one of the co-presenters at the Miwaukee conference. We thank Shoichi for his contribution to the project. Many thanks are due to the reviewers of the paper, whose detailed and critical comments are responsible for its much improved end results. We also thank Maggie Camp for going over multiple versions and making sure that it coheres as a linguistic paper and Yuka Matsugu for her editorial assistance.

speakers choose the language they use). Needless to say, during the initial stages of fieldwork, we are naturally extremely conscious of the socio-cultural factors at work. We pay a great deal of attention to the processes of entering and establishing our status in the language community, finding fluent speakers of the local language to work with, and building personal relationships. However, once we have established a relatively stable working relationship with suitable speakers and are immersed in the linguistic work, the focus of our attention naturally shifts to the form of language itself. Consequently it is rather easy to slip into naïve assumptions about the fieldwork situation. That is, we tend not to be concerned much about the effects of socio-cultural, interpersonal or interactional factors on the production of linguistic forms in fieldwork sessions, expecting fluent speakers to produce their language the same way as they do when they speak among their community members daily. In addition, we tend to view the speakers, though unknowingly and unintentionally, as language robots who can mechanically produce and manipulate any linguistic expression upon request.

In this paper, we would like to draw attention to the fact that language production is sensitive to the contextual factors present in common fieldwork settings to such an extent that they can significantly shape our data gathering and ultimately affect our analysis and documentation. Specifically, we highlight one particular type of contextualization that we regularly experience or, more precisely, inadvertently create while trying to document Ikema.

Ikema is a dialect of Miyako Ryukyuan, a 'definitely endangered' language spoken on Japanese islands near Taiwan (Moseley 2010; Fija et al. 2009). In particular, we will see that, due to the history between the Okinawan and Japanese languages and the specific roles which these languages play in the life of the Ikemas, the very presence of a Japanese researcher with a well-established elevated social status, such as a university professor (or "master," as they are known to locals), may affect not just the production of Ikema but even the so-called native speaker intuition. This is one of the multiple layers of context that constitutes complex fieldwork situations (Dimmendaal 2001; Crowley 2007), and being aware of such contextualization allows us to be more careful in choosing how to gather and interpret data.

2. Background

Miyako, of which Ikema is a dialect, is one of several Ryukyuan languages which are genetically related to Japanese (Japonic). Ryukyuan languages (and their dialects, including Ikema) and Japanese have been in daily contact for the past several decades. People in the Ikema community now speak Japanese more or less

fluently while children are no longer growing up learning their heritage language. Generally speaking, it is rare to find fluent Ikema speakers among the population below the age of 55, and it has been estimated that there are only 2,000 fluent speakers total (Kawada et al. 2009; Iwasaki & Ono 2011).

Currently, Ikema is spoken in three main communities. The oldest Ikema-speaking community is on Ikema Island, a small island north of Miyako Island. Miyako Island is itself located 170 miles southwest of the main Okinawa Island and just 180 miles northeast of Taiwan. Until a bridge was opened between Ikema and Miyako in 1992, Ikema was only accessible by sea. In 1874, some residents were forced to relocate from Ikema Island to the northern part of Miyako, forming a community called Nishihara. The third Ikema-speaking community is Sarahama on Irabu Island, which was established approximately 300 years ago. Irabu Island, located northwest of Miyako Island, is currently only accessible by sea, but a bridge is currently under construction with an opening date set for 2015. Even after a long period of separation, speakers of the three disparate communities still identify themselves as *Ikema Minzoku* 'Ikema Race'. This strong sense of identity is probably shaping the profile of the Ikema dialect being relatively distinct from other varieties spoken on Miyako and nearby islands. Our research team has been conducting fieldwork mostly in Ikema and Nishihara.

Hundreds of islands, including Miyako and adjacent islands located between Taiwan and Kyushu Island of Japan, formed part of the Ryukyu Kingdom, a politically independent nation that no longer exists. Perhaps due to its location, the Kingdom initially established a strong bond with China through trading. Under heavy sociocultural influences from China, Ryukyu developed an original culture, during the Kingdom era, including art, music and religion. This included its regular practice of Confucian ideas, where relationships among people are rather clearly hierarchized, one example of which is the master-disciple relationship highlighted in this paper. The Kingdom, however, was invaded by the Satsuma clan of Kyushu backed by the Japanese government in the early 17th century and officially annexed to Japan in the late 19th century. Under the Japanese government's "reformation," an assimilation policy in all but name, Ryukyuan culture and language were more or less completely suppressed. This was part of Japan's expansion into the Asia-Pacific, where Taiwan (1895) and Korea (1910) were colonized in similar manners.

Perhaps the best-known language-related example of this policy is the infamous *hoogen fuda* 'dialect placard', a device that was employed at school (mostly in elementary schools) in order to discourage the use of the vernacular language, which was designated as *hoogen* "dialect" (Itani 2006; Kondo 2008). At that time, the use of the local languages was strictly forbidden at school. When students slipped into the vernacular language, they were forced to wear a wooden tag with *hoogen fuda* 'dialect placard' or *hoogen shiyoosha* 'dialect user' written on it as part

of their punishment, which sometimes included physical punishments as well. It is not known who initiated this practice or how, but it was already in practice in the early 20th century, and apparently the teachers played major roles in enforcing it.

This language policy, including the utilization of the dialect placard, seems to have firmly instilled inferior views of their vernacular language among the Ryukyuan people. Thus Ikema and Japanese are in clear contrast in terms of their social status: Ikema, without prestige and with social stigma,[1] is used in vernacular contexts, whereas Japanese with its high social prestige is used in formal and official contexts. The language situation in the area can thus be characterized as a typical diglossia with Japanese functioning as H-language and Ikema as L-language (see Ferguson (1959) and Fishman (1967) for characteristics of diglossia). This sociolinguistic stratification has been established and reinforced through formal education (Itani 2006; Kondo 2008; Fija et al. 2009), and has resulted in the particular Japanese-speaking contextualization highlighted in this paper.

Interestingly, the choice of language between Ikema and Japanese seems to be connected more closely to contextual style than social identity. Thus, Ikema speakers choose to speak Japanese simply because it is *contextually* appropriate to do so. Speaking Japanese does not necessarily constitute conflict with or negation of their social identity as Ikema people. This situation is making it difficult to motivate speakers to expand or even keep their use of Ikema as a vernacular language.

It should be noted that, as we spent more time in the area, it became clear that locals actually use three language varieties, Ikema, Standard Japanese, and the local version of Japanese, with varying levels of fluency and frequency in various contexts. It is also becoming apparent that switching among these varieties is a rather common but complex phenomenon controlled by a number of social factors.[2] Below, however, we will focus on the distinction between Ikema and Japanese without distinguishing Standard Japanese from the local version of Japanese.

3. Initial experience

When we first started our work in the Ikema community several years ago, we were met with strong resistance from the islanders toward speaking Ikema. Simply put, the islanders refused to speak Ikema in front of us; we heard very little Ikema in our initial trips. Only after several years of periodic visits to the area

1. The use of Ikema in religious ceremonies is perhaps one exception for this generalization.

2. Needless to say, determining the factors involved in this switching is a project of its own.

have we become more successful in having them produce Ikema for us. Even then, we are finding that the momentum for speaking and thinking in Ikema is fragile. The binding effect of Japanese is incredibly strong: we are becoming increasingly suspicious that speakers' intuition about the grammatical patterns of Ikema is not immune from the influence of Japanese linguistic patterns.

Undoubtedly, this interference is triggered partly as a form of accommodation: the speakers have to speak to complete outsiders, mainlanders who have little linguistic facility in Ikema but speak Japanese, the language of prestige. However, we noticed that the interference problem was less prominent when students interact with them, even though they were equally outsiders and also had little (sometimes even less) linguistic facility in Ikema. It is interesting to point out that even when it was an Anglo-American student who appeared the least likely to be able to understand Ikema, the speakers still went ahead and spoke it in front of her.

4. Contextualization by *shinshii/shiishii* 'master'

As we spend more time in the local communities interacting with islanders on various occasions, it is becoming increasingly clear that our social status is causing a stronger and longer-lasting effect on the speaker. We are university professors from Mainland Japan and North America and, therefore, we socially belong to one of the most exalted types of *shinshii/shiishii* 'master'. This category is perhaps given the highest regard in the Ikema community, where most of the people we deal with, especially the elders, have a rather limited amount of schooling. (Elders only had a handful of years of schooling while the majority of younger speakers completed high school.)

In fact, many factors in our fieldwork settings work to strengthen the formality triggered by our status as *shinshii/shiishii*: our age/gender profiles (male; middle-aged) match the typical image of *shinshii/shiishii*, and our sessions are conducted entirely in Japanese[3] and typically take place at the community centers and houses of prominent members of the community. The consultants evidently take this formality very seriously: although people from the area are apparently rather relaxed about appointments and commitments, we have found our consultants to

3. One might suggest that one step to solve our problem is to conduct sessions in Ikema, which was initially not possible due to our complete lack of command in the language. However, it should be stressed that our communicating in Ikema is extremely unnatural: all the Ikema speakers also speak Japanese, the established medium of communication with outsiders, either from other areas of Okinawa or Mainland Japan. That is, speaking Japanese to us is the most natural, perhaps even automatic, behavior for the islanders.

always be very punctual, responsible and dressed up, and often serve us food, give us gifts, etc. This is quite different from the experiences of locals (those close to us) and students who report that appointments (and even promises) can be easily cancelled and missed.

Thus, our social profile automatically brings out an interactional space that strongly favors the use of Japanese (inadvertent contextualization). The speakers are quick to respond to this contextual setting and switch to Japanese. This naturally influences the type of language which they use to and in front of us. In contrast, students do not seem to trigger the Japanese-speaking frame, at the least not to the same extent, as long as they do not reveal their affiliation to well-known prestigious universities.

5. Episodes from work sessions

One situation highlighting the contextualization effect discussed above was encountered during a recording session in which one of the authors of the current paper participated as a video camera operator. The sequence where a graduate student was eliciting a story from an elder woman in Nishihara reveals the effect the social identity of the researcher has on language choice. The speaker was telling a story where she saw ghost dogs, one of which was the dog she used to have. At the end of the story told in Ikema, the student asked in Ikema where the dog went after she saw him. Interestingly, the elder's answer to this question was not in Ikema, but in Japanese *wakaran moo inakatta* '(I) don't know. (He) was already gone,'[4] even though both the student and the speaker herself had just been using Ikema.

It should be pointed out that this particular student had become quite fluent in Ikema by the time of the recording session; she had the best speaking skills in our project team. In fact, from other interactions between the speaker and the student during the same recording session, it is certain that the speaker knew that she could communicate with the student in Ikema. She could very well have responded to the question in Ikema. Still, the speaker switched to Japanese. Interestingly, after this question and answer exchange, the woman switched back

4. One might expect the polite form *wakarimasen moo imasen deshita* instead of *wakaran moo inakatta* '(I) don't know. (He) was already gone.' It is most likely that Japanese polite forms are not part of the linguistic repertoire of this particular speaker partly because of her age (around 90 years old at the time of recording); she doesn't use polite forms even when she talks to the professors.

to Ikema and continued her story with a focus on her dog, perhaps in response to the student's question.

What is happening here might be accounted for by the often-noted distinction between the real and story worlds (e.g. Clancy 1980). The latter is the world of the story which the speaker just created in Ikema. The former in this context is a recording session for research at the community center. In this 'real world', there were three researchers present, and the whole event was being video recorded by one of the authors of the current paper, who is a *shinshii/shiishii* 'master'. All of these factors were contributing to heighten the formality of the recording setting. Such characteristics of the setting seem to have induced the speaker to switch to Japanese when she was pulled out of the story world to the real world by the student. Thus, the formality of the recording session overrode the encouragement from the conversation partner (i.e. the student) in the speaker's use of Ikema.

Our cases may not be surprising considering that Japanese is known to be overtly responsive to such contextualization[5] yet it needs to be stressed that it took several years for us to start recognizing the pattern. Part of our purpose here is to sensitize other, especially new researchers, to this kind of issue so that it will be a quicker process for them, since it is very likely that switches similar to the ones we report in this paper happen when working on other languages (e.g. Blom & Gumperz (1972) where bilingual speakers were reported to choose which code to speak depending on where they were and what they were discussing).

In addition, we have come across a similar but perhaps more striking situation reflecting elders' views about Ikema and Japanese. There was a recent documentary TV program created by the Okinawa branch of *Nippon Hoosoo Kyookai* (National Broadcasting Company). The nationally broadcast program featured a new daycare facility for elders on Ikema Island. This facility has built its reputation as a successful program by taking care to provide elders with a relaxing, home-like care environment. Use of the heritage language is one of the emphases of the program: daycare staff members are mostly locals and speak to elders in Ikema. What was interesting about the TV footage was that, in a number of scenes, daycare staff members spoke to elders in Ikema (thus subtitled in Japanese) only to get responses from elders in Japanese. Below is one example in which an elderly woman who was living in a facility off Ikema Islands for an extended period of

5. Another way of saying this is that Japanese has resources that speakers use to constitute various situational contexts.

time had just returned to her house because she could now live there while being looked after by the staff at the daycare center. The woman was very emotional:

(1) staff: *hukarasI munu ii* (Ikema)
'Happy, right?'
elder: *hai* (Japanese; polite form)
'Yes'
staff: *yaankai tti ii* (Ikema)
'Coming home, right?'

When spoken to in Ikema by the staff, the elder replied *hai*, the polite affirmative response expression in Japanese.

There is, in fact, something marked about this language choice. Generally speaking, elders are not as fluent in Japanese as younger staff members, whereas the latter are not as fluent in Ikema as the former. Thus, if the language was chosen strictly on the basis of fluency or ease of use, elders should not have chosen Japanese, especially when they were spoken to in Ikema. Furthermore, by following Confucianism which, as we said earlier, is rigorously practiced in this area, it's the younger speaker who should have accommodated to the elder. However, it was actually the elder who accommodated to the younger by speaking in Japanese. What is different between the everyday interaction at the daycare facility and the interactions on the TV program was the context of interaction: the TV shooting affected the context in such a way that the settings apparently became public and formal, thus a Japanese-speaking context for the woman.

In our fieldwork, we have become increasingly aware of the fact that we unknowingly bring a set of contextualization to our work sessions, and that this might affect the type of data we obtain there. Data collection in fieldwork takes various shapes, and fieldwork in descriptive research typically combines different methods in order to build a well-rounded database. Discourse data are useful in capturing the actual use of language, while elicitation is an efficient and effective way to systematically gather forms and structured patterns of the language. It is relatively easy to imagine and in fact see the effect of contextualization in discourse data, as production of discourse is naturally embedded in social and interactional contexts. The context affects not only the content but also choices of language and of linguistic forms. Elicitation, on the other hand, may be considered immune from contextual effects: language forms that are unrelated to each other are elicited out of context, and therefore speakers are assumed to focus more on the mechanical alternation and makeup of linguistic expressions rather than the context of use. Unfortunately, this does not seem to be the case.

6. Elicited data

In this section, we would like to report cases in which elicited data seem to have been influenced by the particular context we bring, i.e. data collection sessions conducted in Japanese by Japanese-speaking university professors. Please note that code switching between Ikema and Japanese appears to be so common in current everyday life for the islanders that some of the grammatical features of Japanese might actually have become accepted as Ikema, especially among younger speakers. Still, we think the particular context of elicitation, an unnatural situation to say the least, promoted the tendency for the speaker to be influenced by Japanese.

Examples below are all taken from our sessions with a single speaker, Mr. G from Ikema Island, who has been one of our main consultants. Mr. G is 59 years old (April 1, 2012) and therefore a relatively young fluent speaker (Iwasaki & Ono 2011). We have had similar experiences with other Ikema speakers, especially Mr. N from Nishihara (mid 60s). We are focusing on data from Mr. G here simply because we have audio recordings of many of work sessions with him, which we have been able to closely study to report on here. We realize that it would be ideal to have illustrative examples from different speakers, but the type of phenomena we are dealing with in this paper are observed haphazardly and are very difficult, if not impossible, to collect systematically. Moreover, peculiarities of the examples that we will report below are so striking even in comparison with the ordinary language use of Mr. G that it seems likely that these peculiarities are caused by contextual effects of working with a *shinshii/shiishii* engaging in an interactionally unusual activity of grammatical elicitation.

We would like to examine examples in three separate areas: pronunciation, verb inflections, and case markers. The wide range of areas where the phenomena could be found suggests the pervasiveness of the potential influence that we are trying to highlight.

6.1 Pronunciation

The first example comes from the area of pronunciation. As we were going over with Mr. G a recorded story, we came across the following utterance:

(2) *gakkoo nu dote no aabi kara*
school of bank of upper.part from
'From the top of the bank at school'

We asked Mr. G for an antonym of the Ikema word *aabi* 'upper part'. Initially Mr. G produced *shita* 'lower part' saying that the expression in Ikema is the same as in Japanese. Interestingly, later in the same session we found him volunteering other

forms *shitanma/shitaara* and *sItanma/sItaara*, stating that *shita* 'bottom' is in fact Japanese and not Ikema.[6]

The form *sIta* in the Ikema words *sItanma/sItaara* appears to be a cognate with the Japanese *shita* 'lower'. The suffixes *-nma* and *-ara* mean 'part/place'. It looks as though the forms *shitanma/shitaara* were innovations based on the native Ikema words *sItanma/sItaara* 'lower part' with the influence of Japanese *shita*.

It is quite possible that the innovated forms *shita(nma/ara)* are in the process of replacing or at least becoming alternative forms of *sIta(nma/ara)* in present-day Ikema, but the Japanese speaking context created by the presence of the researcher may very well have prompted Mr. G to be under a stronger influence of Japanese and to produce the influenced forms *shita* and *shita(nma/ara)* instead of the Ikema version.[7]

6.2 Verb inflections

The example below also came up in elicitation while going over a text. The text was a narrative about an old fishing practice using dynamite. The speaker is explaining how the dynamite was prepared.

(3) *kiigama hii gufugufu ti usunkiutai dara yo*
'(He) was like pushing it in (into a bottle) with a piece of wood'

gufugufu	*ti*	*ttaki-nki*	*usu-nki*
onomatopoeia	like	hit-in	push-in

'hitting and pushing in'

With regard to the verb form *ttaki-nki* 'hit in' we asked Mr. G for the base form of the verb. The form Mr. G volunteered at first was *ttaku* 'hit'. However, the typical form Ikema speakers use to cite a verb is the dependent form ending in *-i*. This is the form that Ikema speakers, including Mr. G, provide almost exclusively when asked for verbs. This is also the form used as the head word form for verbs in word lists and dictionaries that are privately compiled or published by Ikema speakers. Thus, the native intuition calls for the *i*-form as the 'basic' or default form. In that regard, the form Mr. G volunteered was clearly an unusual one. We were puzzled at this form, so we checked other inflectional forms including:

(4) a. *ttatsI bau* 'hitting stick'
b. *ttakii* 'hitting and'

---

6. The symbol 'I' is used to represent a high central vowel.

7. One of the reviewers suggests that this shift belongs to the area of morphology which we understand to mean a switch in an individual form from *sIta* to *shita* 'lower'. The pattern, however, appears to be more general in that a high front vowel is observed replacing a high central vowel in Ikema words.

c. *ttakiiui* 'is hitting'
d. *ttaki* 'Hit!'

Interestingly, after working through different forms and word combinations, Mr. G later volunteered that *ttaku* is in fact Japanese, apparently based on the base form of the Japanese verb 'hit': *tataku*. Over the several years of working with Mr. G up until this point, he had consistently provided us with verbs with the expected ending *-i*; we hadn't really been given them with *-u*. It is therefore unlikely that the use of *u*-ending verb forms is part of a general change due to Japanese influence. Instead, Mr. G's Ikema intuition was likely influenced partly by the rather unusual and unsettling communicative situation of having to go over a verbal paradigm upon request from a *shinshii/shiishii*. It is probably the case that the last time Mr. G had to do something similar was when he was still in school going over Japanese verbal paradigms (this is part of the Japanese school curriculum).

6.3 Case markers

The third illustrative case comes from an elicitation session in which we were examining the use of Ikema case markers. Unlike Japanese, the marking of the subject arguments of clauses in Ikema follows the so-called animacy hierarchy: *ga* marks subjects which are pronouns, proper names, etc. and *nu* non-humans, including animals, things, etc. (Iwasaki & Ono 2010). In Japanese, the subject can be marked with *ga*, but not by *no*, regardless of "animacy."[8] The case forms in Ikema and those in Japanese are cognates. We were checking on the use of *ga* with the subject argument *mayu* 'cat', an animal. Please note that the subject argument *neko* 'cat' in Japanese can be marked with *ga* as in:

(5) *koko ni neko ga kita yo*
here LOC cat NOM came FP
'A cat came here.'

In fact, replacing *ga* with *no* seems to produce an "ungrammatical" sentence in Japanese. In contrast, in the Ikema version of this utterance, *mayu* 'cat' is expected to be accompanied by *nu* because it is an animal, which Mr. G accepted as in:

(6) *umankai mayu nu du ttai doo*
here:to cat SUBJ FOC came FP
'A cat came here.'

8. However, see Ono, Thompson and Suzuki (2000) for an alternative view of the status of Japanese *ga*.

Then, in trying to confirm the animacy-based alternation of case markers, we tried replacing *nu* with *ga* in the above example:

(7) *umankai mayu ga du ttai doo*
here:to cat SUBJ FOC came FP
'A cat came here.'

To our surprise, Mr. G also accepted this even though the use of *ga* on animal subjects had previously been consistently unacceptable to Ikema speakers, including Mr. G himself. We were puzzled but, interestingly, Mr. G later volunteered that the above example actually sounds like Japanese and not Ikema. Again, what may have happened was that Mr. G was influenced by the grammar of Japanese, which may have been foregrounded by a formal setting created by our presence.

Thus we have seen examples in three different structural areas in which elicited data in Ikema may have been influenced by the dominant language, Japanese. That is, the particular contextualization which the researcher brings to linguistic work sessions seems to influence or even override the native speakers' intuition. It has taken several years for us to realize this type of subtle influence on Ikema from Japanese.

7. Discussion

We can see from the above examples that the choice of language or language variety is very sensitive to contextual factors in the Ikema community. It is not easy to tease apart the various socio-pragmatic contextual factors at work in fieldwork, but, considering the way the social relationship and communicative exchange is structured, we suspect that our social identity plays a significant role in shaping the context. Pragmatic contextualization is anchored to discourse context and is relatively easy to detect, because the contextual effects are reflected in variations in the data. However, the contextualization tied to the researcher's social identity is rather tricky to deal with. The social profile, e.g. our social status as *shinshii/shiishii* does not change or 'wear off' quickly. This makes it extremely difficult to recognize the effect of contextually triggered language choices: the affected language use pattern is all we see for an extended period of time, and as a result we could easily be misled into believing that what we see/hear is the normal and complete picture of language use. The only chance for us to realize what is really happening is when we see/hear the data collected by other researchers with social profiles different from our own. In fact, that was how we have come to realize the problem after several years of fieldwork.

Needless to say, this finding itself is not necessarily surprising or new. However, such a context-sensitive language choice has been discussed mostly in connection to sociolinguistic variations and has not been problematized much in reference to fieldwork situations. This is likely due, in part, to the belief that the phenomenon is exceptional and relevant only in a limited domain of language use.

However, the reality is that language switching is pervasive and very much a part of 'natural' language life, at least in the community we work in. Thus, for Ikema speakers, the switching between Ikema and Japanese is a matter of style choice rather than language choice (a structural system of communication): that is, the two language varieties have become 'ways of speaking' associated with particular domains in everyday life. Note that this situation is not something that has happened recently in connection to the endangerment of Ikema. Minimally, Ikema and Japanese have been in daily contact in the lives of currently living community members, which equates to approximately the past 100 years.

It is interesting to point out that this kind of active switching might actually be observed more in communities like Ikema where the heritage language still has relatively active usage domains. Language choice is clearly associated with different parts of everyday contexts, and language switching is a natural reaction to contextual change. Perhaps the situation is different in more severely endangered language communities. In such communities the vernacular language is completely removed from everyday life, and consequently language choice could be more of a matter of conscious decisions.

If we were to seriously consider the fact that language switching is a common, natural part of language use even in fieldwork encounters, we need to adopt a different attitude toward the switching and mixing of language use and embrace it as part of the natural state of the language. Given that language choice is constantly affected by different contextual factors to varying degrees, it may be the case that, by insisting that the Ikema speakers speak to us in their vernacular language, we are forcing them to use their language in a very unnatural way, i.e. in a situation where they don't normally use or feel uncomfortable using the heritage language. As Ikema has been perceived as a 'non-standard' speech variety, it is commonly associated with very informal, in-group situations. In the fieldwork setting, however, speakers interact with outsiders engaging in unfamiliar (and most likely unnatural) activities. If we add this to the presence of researchers from Mainland Japan and North America who carry a high social status, the resulting environment will most certainly be the worst possible match with the use of Ikema. Asking speakers to speak in Ikema regardless of the context, then, would be as unnatural and outrageous as asking someone to speak rudely to you for the

sake of language data collection. Thus, when we ask Ikema speakers to speak in their vernacular language, we put the speakers in a double bind situation where they feel pressured to honor the request to speak in Ikema but at the same time feel resistance to speaking in Ikema out of respect to us. It may be dangerously naïve to expect language produced in such conditions to be the perfect representative of use of the target language.

An even more important problem of the language choice is its effect on grammatical intuition. We tend to think that language choice is a matter of superficial language encoding: that is, as long as the speaker is using Ikema linguistic expressions, s/he is operating in the 'Ikema mode' from the top to the bottom. However, as our examples showed, there seem to be cases where the speaker's structural intuition (or framework of structural understanding) also shifts according to the context. Even when the speaker is using one language, s/he may be thinking about the patterns in terms of the other language's structural frame: thus, even when the speaker is using Ikema words, the way s/he understands and thinks about the patterns could very much be framed in the organization of the Japanese grammar. This can happen particularly when the speaker is educated in the other language, as in the case of Ikema speakers who were mostly educated in Japanese. It is also important to realize that being asked about grammatical intuitions and structural patterns is not something that happens regularly in ordinary life. The closest analogue to the fieldwork environment that our speakers might have is probably the classroom where they studied about Japanese grammar. This explains why Ikema speakers might bring in the Japanese grammatical context to discussion of Ikema expressions and forms.

8. Conclusions

It should be clear from the above discussion that inadvertent contextualization is a source of serious concern for field linguists. Given that language is fundamentally a social phenomenon, the contextual shaping of usage pattern is very much a natural part of language. Although researchers tend to forget this fact, we are very much part of the 'context' that affects the way language is used. This is very unfortunate and troubling for researchers. We cannot control or eliminate the effect of who we are, and that means it is not possible for us to avoid, eliminate, or even control completely the problem or the artifacts in the data introduced by our presence in fieldwork situations. Needless to say, it is not our intention to question the validity and usefulness of field linguistic research, but to point out the need for awareness of the issue and the danger of naively assuming that

linguistic data collected through fieldwork is fundamentally natural and socially neutral. In fact, being aware of the potential impact of contextualization in gathering and interpreting field data is probably the best countermeasure against the contextualization problem.

Another effective way to mitigate, if not to solve, the problematic effect of contextualization is to diversify the context of language data collection by working in a team consisting of members with different social profiles in terms of age and gender, among other factors. The importance of collecting a wide variety of data has long been discussed by researchers from the perspective of coverage of discourse types. As pointed out in recent studies, grammar is different for different genres, goals, contexts, etc. (Hopper 1998; Iwasaki 2006). We, therefore, need to capture language use in as wide a range of discourse types as possible in order to obtain a well-rounded picture of the language. What our discussion showed is that, in addition to the traditionally recognized contextual factors, it is important to pay attention to the identities of the researchers as a contextual factor to control for.

The strategy that we have found particularly useful in tackling the problem of inadvertent contextualization is to involve members of the local community in data gathering sessions. In fact, part of what made us notice the problem was examining recordings made by one of our Ikema-speaking collaborators without our involvement. Her recordings were noticeably more natural and free of Japanese than those we made precisely because they were made without the layer of context which would have promoted the use of Japanese. Similarly, sometime after one of our project members (another middle aged male professor) recorded stories told by an elder couple, we asked Mr. G. to visit the same couple to record them. It turned out that the couple produced some of the same stories, which were expectedly much more freewheeling and involving much less Japanese than the version they initially told us.

We go out to the field in order to obtain a firsthand sense of the natural state of the language. We hope to have shown in this paper that this 'natural state of the language' is very fragile and that our presence can have a powerful and disruptive effect. This is not a trivial problem, especially when it can affect the structural framing of speaker's understanding of language patterns, i.e. 'grammatical intuition'. The tricky aspect is that the mechanism of the inadvertent contextualization is, in a way, part of natural language use, and as such, there is no easy way around this problem when we engage in fieldwork. We can never be socially neutral or invisible. We are socially real people, even in linguistic elicitation where we are supposed to be focusing on just the structural aspect of language. Recognizing this fact is a significant first step forward.

References

Blom, Jan-Petter & Gumperz, John J. 1972. Social meaning in linguistic structures: Code switching in northern Norway. In *Directions in Sociolinguistics*, John J. Gumperz & Dell Hymes (eds), 407–434. New York NY: Holt, Rinehart, and Winston.

Clancy, Patricia. 1980. Referential choice in English and Japanese narrative discourse. In *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*, Wallace Chafe (ed.), 127–202. Norwood NJ: Ablex.

Crowley, Terry. 2007. *Field Linguistics: A Beginner's Guide*. Oxford: OUP.

Dimmendaal, Gerrit J. 2001. Places and people: Field sites and informants. In *Linguistic Fieldwork*, Paul Newman & Martha Ratliff (eds), 55–75. Cambridge: CUP.

Ferguson, Charles. 1959. Diglossia. *Word* 15: 325–340.

Fija, Bairon, Brenzinger, Matthias & Heinrich, Patrick. 2009. The Ryukyus and the new, but endangered, languages of Japan. *The Asia-Pacific Journal: Japan Focus* 3138.

Fishman, Joshua. 1967. Bilingualism with and without diglossia; Diglossia with and without bilingualism. *Journal of Social Issues* 23(2): 29–38.

Hopper, Paul J. 1998. Emergent grammar. In *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, Michael Tomasello (ed.), 155–175. Mahwah NJ: Lawrence Erlbaum Associates.

Itani, Yasuhiko. 2006. *Okinawa no hoogenfuda* (*Dialect Placards in Okinawa*). Naha: Borderink.

Iwasaki, Shoichi. 2006. The 'multiple-grammar' hypothesis: Case studies of Japanese and English. Invited address at Association of English Linguistics, University of Tokyo.

Iwasaki, Shoichi & Ono, Tsuyoshi. 2011. Ikema Ryukyuan: Investigating past experience and the current state through life narratives. *Japanese/Korean Linguistics* 19: 351–364.

Iwasaki, Shoichi & Ono, Tsuyoshi. 2010. Differential subject marking in the Ikema dialect of Miyako. Paper presented at the 9th High Desert Linguistics Society Conference, University of New Mexico.

Kawada, Takuya; Hayashi, Yuka; Iwasaki, Shoichi & Ono, Tsuyoshi. 2009. Ryukyugo Miyako Ikema hoogen ni okeru *mmya* (The Discourse function of *mmya* in the Ikema dialect of the Ryukyuan language). In *Ryukyu Okinawa no Bunka Tayoosei: Seitai, Ninchi, Gengo* (*The Cultural Diversity of the Ryukyus/Okinawa: Ecology, Cognition, Language*), R. Shibasaki (ed.), 111–130. Itoman, Okinawa: Bunshin Insatsu.

Kondo, Kenichiro (ed.) 2008. *Hoogenfuda* (*Dialect Placards*). Tokyo: Shakaihyooronsha.

Moseley, Christopher (ed.). 2010. *Atlas of the World's Languages in Danger*, 3rd edn. Paris: UNESCO Publishing. ⟨http://www.unesco.org/culture/languages-atlas/en/atlasmap.html⟩

Ono, Tsuyoshi, Thompson, Sandra A. & Suzuki, Ryoko. 2000. The pragmatic nature of the so-called subject marker *ga* in Japanese: Evidence from conversation. *Discourse Studies* 2: 55–84.

Internal and external calls to immigrant language promotion

Evaluating the research approach in two cases of community-engaged linguistic research in Eastern North Carolina*

Ricard Viñas-de-Puig
East Carolina University

This paper explores the research methodology of the documentation and promotion of two immigrant indigenous languages in Eastern North Carolina, by contrasting it to a successful participatory research approach observed in an ongoing project in Nicaragua. When contrasting immigrant and *in-situ* settings, it becomes apparent that the different priorities of both communities do not allow for the implementation of an identical participatory research approach. To overcome these differences, the research approach needs to be adapted to the needs of the immigrant community, while still maintaining a strong community-centered focus that enables an egalitarian relationship between researcher and members of the speaking community, and that provides services and materials that first and foremost benefit such community.

1. Introduction

In recent years, in the linguistic literature there has been significant interest in the ethical methodology employed when doing language documentation and revitalization research (Bach 1995; Christie et al. 2000; Benedicto et al. 2002; Rice 2004;

* I would like to express my sincere gratitude to the Tzotzil and Hñähñu speakers of the immigrant community in Eastern North Carolina for allowing me to collaborate with them in these projects aiming at the promotion and preservation of their languages in circumstances that are not always the easiest. I also want to acknowledge their hospitality and their willingness to share their language.

This paper has also benefited from extremely valuable comments from different reviewers, who also deserve my gratitude. Any errors remain my own.

among many others). However, most of this methodological research has dealt with efforts carried out in the preservation and documentation of languages in the original communities where these languages are used (Christie et al. 2000; Benedicto & Mayangna Yulbarangyang Balna 2007; Czaykowska-Higgins 2009; Viñas-de-Puig et al. 2012; among others), but very little, if any, attention has been paid to the promotion of immigrant languages in the US (and elsewhere). In many cases, these languages (and the research involving them) have remained 'invisible' not only to the general public but also to the community where these languages are present.

This paper attempts to partially fill this void by presenting two cases of community-engaged research implemented in Eastern North Carolina since 2009, and by exploring how these two cases make us revisit some methodological aspects of participatory research. Thus, the importance of the paper relies not only on the fact that it exposes the challenges of a participatory research methodology when working with immigrant indigenous languages,[1] but also, and maybe more importantly, because it sheds light on a linguistic situation that is too often unknown, or even ignored.

1.1 Goals and summary

The main goals of this paper are twofold. The first objective is to present an overview of the presence of immigrant indigenous (or underrepresented) languages in the US, and especially in North Carolina, and to highlight some of the efforts underway to promote the use and preservation of these languages by first and second generation immigrants. The second, and most important, goal of the paper is to analyze the implementation of a participatory research methodology in these documentation and preservation efforts, and to discuss the challenges observed in this very specific type of linguistic research.

With these goals in mind, this paper is targeted not only to external researchers (e.g. linguists) working with these languages, but also to social workers, community leaders, and other people involved in those communities in the US, and elsewhere, where immigrant underrepresented (or even endangered) languages are present.

---

1. As noted by one anonymous reviewer, the term 'immigrant indigenous language' is a contradiction, since an immigrant language is, by definition, not indigenous. Despite of this contradiction, I decide to keep this term (for a lack of a better one), since this paper deals with languages that are indigenous to different regions of the Americas (considering this as a macrolinguistic area), thus excluding those languages of European (or other) origin.

1.2 Paper outline

This paper is organized as follows. Section 2 presents a basic overview of the recent immigration patterns in the US, and how these affect the presence of indigenous (or underrepresented) languages in the US, and especially in Eastern North Carolina. In this section, an overview of the sociolinguistic situation of these immigrant indigenous languages is also discussed. Section 3 examines some of the community-centered research methodologies that have been discussed in recent linguistic literature, with an emphasis on Participatory Action Research. The last subsection in Section 3 presents a successful implementation of such a Participatory Action Research approach in an ongoing project in Nicaragua. Section 4 introduces two ongoing projects in the promotion of underrepresented immigrant languages (i.e. Tzotzil and Hñähñu) in Eastern North Carolina, while Section 5 is a discussion of how these projects present challenges that force a reshaping of the research methodology to meet both the projects' objectives and the communities' needs. Finally, in Section 6 I summarize the main points of the paper and extract some conclusions that might have a relevant impact on the field.

2. Immigration and Mesoamerican indigenous languages

2.1 Concentration of immigrant population in the US

It is a widely known fact that in recent years Latino (or Hispanic) immigration has been on the rise in the US. In many instances, these immigrant groups present a tendency to concentrate in those regions where there are higher working opportunities, as pointed out in Fox (2004). Although there are higher concentrations of immigrant populations in Florida, Texas, or California (Census Bureau 2010), North Carolina has experienced a significant growth in the arrival of immigrant population in the last decade: in the year 2000, the Hispanic population constituted a 3.7% of the total population of the state (Census Bureau 2000), whereas in the year 2010 the Hispanic population had more than doubled, reaching 8.4% of the total population (Census Bureau 2010).

2.2 Arrival of 'invisible' languages

The arrival of these new immigrants, and contrary to popular belief, leads to the arrival of languages other than Spanish. However, and despite some efforts in the identification, promotion, and documentation of these languages (Burke 2002; Fox 2004; among others), still very little is known about their presence, and even less is done to serve the speakers of these languages.

In North Carolina, recent works reveal the presence of some American indigenous languages among the Hispanic immigrant population, although most of these studies have focused on areas that do not include Eastern North Carolina (Fink 2003; Pick et al. 2011). Yet, an also recent survey by Ortega (2010) indicates that indigenous languages are also present in this region of North Carolina; the existence of these languages in this area's immigrant community has led to two independent, but related, projects, which are the object of this paper. Before detailing the research approach and the efforts put forth to strengthen the use of these languages, in the next subsection, I present a general view of the sociolinguistic situation among the (Hispanic) immigrant population in Eastern North Carolina.

2.3 Sociolinguistic situation of immigrant languages in Eastern North Carolina

As indicated above, with the arrival of new immigration population of Latin-American origin, languages other than the majority language (i.e. Spanish) have also arrived in the US. Eastern North Carolina is no exception to this, as made evident in a study by Ortega (2010). In her survey, based in the town of Tarboro, NC, Ortega (2010) presents data revealing the existence of seven different Mesoamerican languages (Mixtec, Q'echi, Popti, Mam, Jakaltek, Q'anjob'al, and Akatek) in a small immigrant community. However, and despite these data, most of these languages are invisible to the general population, and even to members of the same immigrant group, which also includes many monolingual Spanish speakers.

A more general overview of the sociolinguistic situation among the immigrant population in Eastern North Carolina reveals some basic trends. As noted by Ortega (2010) and different field observations (carried out between 2009 and 2011), in those Hispanic communities with presence of immigrant indigenous languages, these languages are in a clear diglossic situation with respect to Spanish and, to a lesser extent, English. The fact that in these multilingual communities Spanish is the lingua franca, which is used in public interactions and in church services, while English is in the vast majority of cases the language of schooling, restricts the use of these other languages to household use (or to a few rare social interactions with other people who speak the same language). It is worth noting that, as Ortega (2010) points out, in some cases there is more than one immigrant indigenous language in a given community; this linguistic heterogeneity favors the use of Spanish as the common language.

Considering the restricted use of these languages in the social sphere, it is not surprising to observe that the younger generations (i.e. children who have been born in the US or have spent most of their life in the US) resort to Spanish and

English as their languages of communication, both in and outside their home. In a few cases, there might be individual cases of children who might be able to understand their parents' first language, but their production is limited.

Another important factor has an impact on the linguistic situation of these immigrant indigenous languages. Eastern North Carolina is a significantly rural area, in which agriculture has a strong impact on the local economy. Therefore, many of the members of the Hispanic immigrant population work in the field (*los files*), performing different tasks and moving to different areas depending on the growing or picking seasons. As a result, some of the Hispanic communities observed in Eastern North Carolina are not only immigrant, but also migrant, and members of these groups move to different towns (or even states) during different times of the year. This has an impact on the linguistic situation of some of the languages that are object of this paper: the presence of an immigrant indigenous language on a given area might suddenly change depending on the season, resulting in a individuals or whole families moving elsewhere due to job related issues.

The arrival of new immigrant population in this area (as noted in the previous subsection), with the subsequent (possible) presence of immigrant indigenous languages, also has an effect, albeit less important, in their sociolinguistic situation. Although it is true that, in gross numbers, there is higher presence of speakers of these languages, the social presence of these languages remains extremely limited: in each community the speakers tend to be in the minority, thus having to resort to Spanish as the language of social use.

All of these factors combined make the situation of these languages precarious at best. Although the immigrants who are L1 speakers of an immigrant indigenous language value it as an important part of their identity, the social, economic, and cultural environment disfavor their using the language and passing it to their children. In some cases, these languages become in danger of disappearing in this immigrant setting, although the situation of the language in the community of origin might not be as negative. This is indeed the case of Tzotzil and Hñähñu, the two languages that are the object of this paper. According to Lewis et al. (2013), both languages are not faced with immediate danger of disappearing in their communities of origin. Tzotzil, an indigenous language of Mexico (mostly spoken in the state of Chiapas) that belongs to the Tzeltalan subgroup of the Mayan linguistic family, is considered to be a threatened language, with over 230,000 speakers, although children are still acquiring it as L1. On the other hand, Hñähñu (or Mezquital Otomí) is a variety of Otomí, an indigenous (macro-)language of Mexico spoken in different states by approximately 100,000 and considered to be in a developing stage, with a standardized and vigorous use. The Hñähñu variety is mostly spoken in the Valle del Mezquital, in the state of Hidalgo, by some 10,000 speakers (Lewis 2009). These situations in origin contrast with what is observed in

the immigrant communities in Eastern North Carolina; in these latter areas, these languages have a very restricted use, have little or no community support, and are not passed on to younger generations.

Subsequently, there is a clear need to develop programs to strengthen the presence of these languages and to promote their presence and use among children and young adults; these were the motivations behind two small projects to foster the use of two immigrant indigenous languages in two immigrant communities in Eastern North Carolina. But before presenting in more detail these projects on the promotion of these languages, in the following section I present a brief sketch and a positive example of participatory linguistic fieldwork. This was the research approach initially conceived when starting the two linguistic promotion efforts presented here.

3. Ethical issues in language preservation efforts

As mentioned in Section 1 above, in recent years there have been a significant number of publications dealing with ethical issues when conducting linguistic research in the field (Bach 1995; Christie et al. 2000; Benedicto et al. 2002; Rice 2004; Benedicto & Mayangna Yulbarangyang Balna 2007; Czaykowska-Higgins 2009; Viñas-de-Puig et al. 2012; among others). In this section, I outline some of the issues that have led to this introspective look at the research methodology, and I pay close attention to Participatory Action Research, one of the research models used in fieldwork in linguistics and other disciplines (e.g. Whyte et al. 1989; Cornwall & Jewkes 1995; Hagey 1997). In the third and final subsection, I briefly discuss an ongoing, successful implementation of the Participatory Action Research approach in a collaborative project with the Mayangna Yulbarangyang Balna, a team of Mayangna linguists based in Rosita, Nicaragua.

3.1 Some problems in linguistic fieldwork

In traditional fieldwork in linguistics (and other social sciences), the research was (and often still is) coupled with a situation of imbalance of power (or the ability to make the crucial decisions regarding the inception, goals, progress, and outcomes of any given research project) between the external researcher(s) and the members of the (speaking) community: in these traditional research situations, the researcher is an expert in the field of study and, as such, is considered as the holder of knowledge; the speakers are considered as research subjects, who become necessary since they can provide the needed linguistic data. From this

position of power, it is the external researcher who creates a need (for knowledge) and contacts a given (speaking) community to gather the information that would fulfill the researcher's need, with little or no implication by the members of the community. In this type of research, which has sometimes been labeled as 'fly-in-fly-out' research (see Eggleston 2012, and references therein for additional information on this type of research), the (speaking) community not only does not participate in the decision-making process, but also usually receives no benefit from the research being carried out. Since the researcher clearly benefits from the exchange (e.g. with the publication of the research carried out using the linguistic data in academic journals), the knowledge gained from the community contributes to widen the knowledge gap; we can easily conclude, then, that this type of research is based on, and promotes, a relationship of power imbalance. In other words, the result of this type of research in may cases does not consider the interests or needs of the members of the speaking community; their participation is limited to provide linguistic data, with no (material or abstract) benefit for them as individuals or as part of a linguistic community.

However, in recent years, many have been the scholars who have advocated for establishing, and many have actually implemented, a more ethical and balanced relationship with the (speaking) community (Bach 1995; Christie et al. 2000; Benedicto et al. 2002; Rice 2004; Benedicto & Mayangna Yulbarangyang Balna 2007; Czaykowska-Higgins 2009; Eggleston 2012; Viñas-de-Puig et al. 2012; among others). One of such approaches, used in different disciplines within the social sciences, is Participatory Action Research. In the following subsection, I present the basic guidelines of such an approach, detailing its different components, goals, and expected end-products.

3.2 Participatory Action Research

Participatory Action Research is a research approach used in different disciplines that advocates for an egalitarian relationship between the external researcher(s) and the members of the (speaking) community, by establishing a relationship that fosters a rebalance of power between both parties involved. The basic proposal of this research model is summarized in two main goals: i. there has to be an active participation of the members of the speaking community, so that they become agents in the discovery process; and ii. there has to be a (self-) empowerment of the community of speakers, validating all 'knowledges' and balancing the power and role dynamics between the external researcher(s) and the members of the community. This resulting egalitarian relationship between all the people involved in the research process is based on three components, outlined in Figure 1.

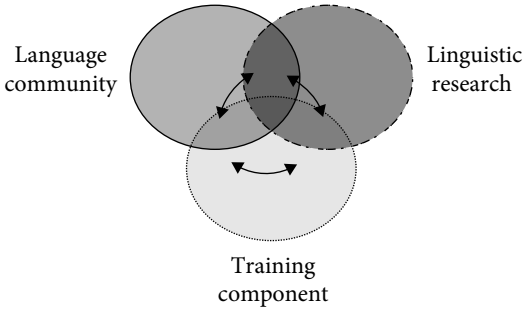

Figure 1. Participatory Action Research PAR components. (adapted from Benedicto & Mayangna Yulbarangyang Balna 2007)

As seen in Figure 1, the Participatory Action Research model is based on three interrelated components aiming at creating a relationship of equal power. The most important component, without which research (and the subsequent implementation of the Participatory Action Research model) is impossible, is the (speaking) community: the community not only provides the goal of this research model (i.e. rebalancing of power), but it also is where language (or whichever object of research) is present.

The second component is research. This is the facet that lies behind the establishment of contact between the community and the researcher(s); in linguistic fieldwork, usually there is interest in a language as a whole, or in a given aspect of a language. However, it is worth noting that in this type of approach the contact can come from either side: as discussed in Section 4 below, the examples of a collaboration that are the object of this paper were triggered both internally and externally to the community.

Training is the third and final component of the Participatory Action Research model,[2] the one that ties the other two components of the research together. This facet ultimately allows for creating a balance of power between all the parties involved and allows everybody to share their knowledge: the external researcher(s) learn(s) from and about the community's knowledge; the members of the community gain the theoretical and practical expertise from the external researcher(s). Since both parties eventually share these different types of 'knowledges', they can informatively contribute to the decision-making process and, consequently, a balance of power is obtained. Crucially, in an ideal implementation of the Participatory Action Research model, the result is that the members of the community

2. For more detailed information on the training facet of the Participatory Action Research model, see Viñas-de-Puig et al. (2012) and the references included there.

become researchers who not only hold the knowledge about their language but also the technical and theoretical knowledge on what to do the research on, and how to conduct such research.

Finally, and following the general principles of the Participatory Action Research model described above, one of the most important goals of its implementation, resulting from the establishment of an egalitarian relationship between all parties, is the creation of products and materials that meet the community's needs. In many cases, these materials not only serve to satisfy the needs of the community, but they are the base of the linguistic research carried out by the external researcher(s). Considering all these goals and objectives, it is worth noting that in this type of research model, although a research outcome (theoretical or otherwise) is still important, this is not the main objective pursued; instead, it becomes a subsidiary result.

3.3 A successful Participatory Action Research project: Mayangna Yulbarangyang Balna

As already mentioned, in recent times several linguists have implemented this type of participatory research in their studies of indigenous and endangered languages. In this subsection, I present an overview of such a project, whose positive outcome demonstrates the validity of this research methodology in language revitalization and promotion projects carried out in the place of origin of the linguistic community. In Section 5 I use this successful example to measure the validity of implementing such a research approach for the promotion of American indigenous languages in US immigration settings, considering two projects in Eastern North Carolina.

Since the mid 1990's, the Mayangna Yulbarangyang Balna, a team of local Mayangna linguists that collaborates with a group of external researchers based at Purdue University, has been working on the documentation and creation of materials on this Misumalpan language. The team was conceived when members of the speaking community contacted external researchers (who were already working on promotion and documentation efforts of another Misumalpan variety) to also work on their language. Since then, the resulting team of community members and external researchers (led by Dr. Elena Benedicto from Purdue University) has been collaborating, documenting the language, presenting research on different linguistic aspects of the language, and creating materials that not only help in the preservation of the language but that are also relevant for the speaking community. These collaborations result in the organization of workshops held twice a year (usually in the months of July and December), when all the participants in the project gather to evaluate the goals and progress of the research, and to hold

linguistic and technology training sessions. During the rest of the year, the members of the Mayangna Yulbarangyang Balna work autonomously on the research projects.

One of the crucial aspects of this example of implementation of a Participatory Action Research approach is the obvious self-empowerment of the members of the community who have also become members of the Mayangna Yulbarangyang Balna. Thanks to the training component of the Participatory Action Research model applied, the Mayangna Yulbarangyang Balna members are now linguists in their own right; some of these members have defended (in Mayangna) their MA theses at URACCAN (Universidad de las Regiones Autónomas de la Costa Caribe Nicaragüense, the local university) resulting from the research conducted, and have presented the results of their investigations at professional linguistics conferences. In sum, the members of the Mayangna Yulbarangyang Balna manifest the results of the self-empowering stemming from a Participatory Action Research model.[3]

Since the initial stages of this two-way project, the members of the Mayangna Yulbarangyang Balna have collaborated with the external researchers to develop materials that are both relevant for the speaking community and the linguistic field. This has resulted in the publication of children books, women's stories, and a monolingual grammar and a monolingual dictionary that can be used in the local bilingual education program.

Considering that the main goals of the Participatory Action Research approach are satisfied in the work carried out by the Mayangna Yulbarangyang Balna (i.e. active participation by the members of the speaking community, and (self-)empowerment of the community of speakers), and that (most of) the materials created are meant for the benefit of the speaking community, we can affirm with a high degree of confidence that this ongoing project represents a successful example of a participatory research approach.

4. Two projects in Eastern North Carolina

As indicated in Section 2, with the arrival of a new immigration population of Latin-American origin, languages other than the majority language (i.e. Spanish) have also arrived in the US. It becomes apparent that the speakers of these languages have little means to preserve their language within their immediate social

3. For a more detailed account of the achievements of (and problems observed with) the Mayangna Yulbarangyang Balna, see Viñas-de-Puig et al. (2012) and Benedicto et al. (2007).

group and to pass it along to younger generations, who resort to Spanish and English as their languages of primary use.

4.1 Two calls, a shared objective

In an attempt to document and provide materials to promote the current and future use (among younger generations) of some of these languages, two independent but related projects have been carried out in Eastern North Carolina since the fall of 2009. Although the main goals (to be discussed in Section 4.2.) and research methodology are the same in both instances, these projects originated under different circumstances.

In the first case, an adult female Tzotzil speaker who is a member of the immigrant community in Eastern North Carolina contacted two faculty members working at East Carolina University to explore the possibility of doing some work with, or on her language. After some initial meetings, it was agreed that this community-internal call (which bears some similarities with the call made by the members of the Mayangna community that led to the establishment of the Mayangna Yulbarangyang Balna outlined in Section 3.3.) would result in a project to create materials so that the language could be passed to the next generation in this immigrant community.

In the second instance, the faculty member who was already working on the Tzotzil project contacted a family of Hñähñu speakers residing in Greenville, NC. The members of this speaking community were mostly adult females; although there are adult males and children in the family, the members who engaged in collaboration with the external researcher are all women.

Although there was no overt interaction between the two groups, this external call to the speaking community resulted in a similar project: the aim of such collaboration was to develop materials so that the descendants of this Hñähñu speaking community had access to written documents in the language.

4.2 Goals of the projects

As stated above, the main goal of both projects was not only to document the languages (and some of its syntactic features, according to the external researcher's interests), but most importantly to develop materials that could be used by the community. Following the basic guidelines of a Participatory Action Research approach, the initial idea was to collaboratively create a series of documents that would be, first and foremost, beneficial to the speaking community, and which could also serve as data for possible linguistic analysis. Therefore, in both projects, the members of the community and the external researcher set as feasible objectives the creation of the following materials: i. a basic multilingual visual dictionary,

that would serve as a teaching tool for children to help them in the active or passive acquisition of the language; ii. a collection of traditional folktales, to build a small repository of oral stories not easily accessible in an immigration setting; iii. a collection of traditional crafts, aiming at building both linguistic and traditional knowledge among the members of the immigrant community; and iv. a series of medical glossaries, that would help the members of the speaking community have a useful tool to better interact with health care professionals. (It is worth noting that this last document would also be one of the primary goals of an international research project on health care service disparities, involving researchers from different institutions in the US, Mexico, and Europe.)

Apart from the creation of end-products that are to be benefited from by the members of the speaking community, these projects also attempt to meet the primary goals of a participatory research approach, as outlined in Section 3.2. above. Consequently, the interaction between the members of the speaking community and the external researcher was conceived from the beginning in a way that would ensure maintaining a balance of power between all the participating parties, and in which the decision-making and methodologies are always shared. To reach such a goal, the initial objective was to create training workshops that would foster sharing and exchanging of the different 'knowledges' brought to the table by the different members of the team.

5. Reshaping the research approach

In the previous section, I outlined the major goals that were established by the different members of the community and the external researcher in an attempt to elaborate promotion materials for both Tzotzil and Hñähñu in their different (immigrant) communities in Eastern North Carolina. In this section, I evaluate the working methodology used in both projects, paying special attention to the challenges that have arisen (especially visible when compared to an *in-situ* research project, like the one established with the Mayangna Yulbarangyang Balna), and how these challenges force the goals of the project in North Carolina to be reviewed and, ultimately, modified.

5.1 Challenges

As mentioned in Section 3.3., the project concerning the Mayangna Yulbarangyang Balna serves as a good example of a successful implementation of a Participatory Action Research approach. Therefore, this project in Nicaragua is a good yardstick against which we can measure the success of a participatory research

project, like the ones established for the promotion of Tzotzil and Hñähñu among the respective immigrant communities in Eastern North Carolina.

When revisiting the success of the Mayangna Yulbarangyang Balna, some attention should be paid to some of the external factors that, although not completely essential in and of themselves to produce successful results, have clearly favored a positive outcome of the project. First, in a situation of linguistic field research *in-situ* (i.e. where the language is originally spoken), we observe a shared linguistic competence: the members of the community who are involved in the project have the same competence in their first and second (and third) languages, which results in working linguistic homogeneity (in the case of the Mayangna Yulbarangyang Balna, all the members speak Mayangna, Miskito, and Spanish).

Second, in a situation of linguistic fieldwork in the place of origin like the one experienced in Nicaragua, preservation of the language and linguistic documentation is an important priority shared by most, if not all, of the members of the community; there is a tacit consensus in the community that a clear effort is needed to preserve and promote the language.

And finally, the members of a team in a situation of *in-situ* linguistic fieldwork also share similar literacy levels: all the members of the Mayangna Yulbarangyang Balna are not only literate in the different languages they speak, but they also are able to use the technological materials at their disposal (laptop computers, audio recorders, video cameras, etc.), which enables them to be involved in all the stages of linguistic research.

The situation in an immigration setting like the one observed in the two projects in Eastern North Carolina reveals clear differences that contrast with the situation described for the ongoing project in Nicaragua. The most important difference observed between an *in-situ* and the immigrant communities is that of language: in both projects in Eastern North Carolina, the languages that originated the study (either Tzotzil or Hñähñu) are not the only languages in either community, since Spanish is the most predominant language in this type of immigration setting. In fact, Spanish is the language passed on to the younger generations, even in those communities (like the one in which Hñähñu is present) where the immigrant language is spoken by most of the members of the community; this is even more so in those immigrant groups in which the immigrant indigenous language is only spoken (and seldom used) by a few members of the group.

Secondly, both immigrant communities in North Carolina showed a different priority ranking from the one observed in Nicaragua: although in all the cases community members value the importance of language and the need for promotion, documentation, and revitalization efforts, in the immigrant communities in Eastern North Carolina, the economic and social situation leads to a clear prioritization of the need to find ways to cover their food and living expenses; in

many cases, working meetings with the community members who were willing to collaborate have had to be cancelled because of external factors. Also, when considering specifically linguistic priorities, it is worth noting that the members of the immigrant community, in an attempt to improve their standard of living in the US, overtly manifest an interest (and a need) to enhance their English skills; however, this need does not enter into conflict with their interest in preserving and promoting the use of their indigenous language.

Finally, the last significant difference observed between a community of origin and the immigrant communities has to do with literacy levels: in the cases observed in Eastern North Carolina, the members of the community had benefited from different access to education, which leads to different levels of literacy competence. The latter fact not only differentiates the team of Mayangna linguists with the Tzotzil and Hñähñu communities in Eastern North Carolina, but crucially it has an impact on the implementation of a participatory research methodology.

A general overview of differences in priorities and skills between the two populations is presented in Table 1.

Table 1. Differences in priorities and skills among participants in origin and immigrant settings

	In Nicaragua	In Eastern NC
Shared linguistic competence	✓	×
Shared community priority	✓	×
Shared levels of speaker literacy	✓	×

5.2 A new research model: From Participatory Action Research to engaged research

Before going into the differences between the originally conceived participatory methodology and the one actually put into practice, a definition of community-engaged research should be provided. In this paper, I understand as community-engaged research that type of research that does not play the center of attention in the research goals, but rather it focuses on the community's needs and provides a result that is clearly beneficial to the members of such community, while at the same time maintaining a balance in the decision-making power dynamics. This is consequently a more general concept than Participatory Action Research, since not all community-engaged research models have to consist of the different components (and resulting relationships) described in Section 3.2.

With this new definition in hand, we can now analyze the implications drawn from differences observed between the immigrant communities in

Nicaragua and North Carolina. These differences force us to revisit the research approach implemented. Since the members of the immigrant communities do not all share a common language, a high priority for linguistic documentation and promotion, and an equal set of literacy competences, the use of a model that requires a homogenous use of all of these (i.e. Participatory Action Research) is no longer viable. However, a focus on a power-balancing project can still be in place.

Once all the challenges in immigrant communities described in the previous section were identified, there was a shift towards a new model that would still maintain a strong community-centered focus. The teams that worked towards the promotion of Tzotzil and Hñähñu, in meetings with the external researcher, (tacitly) changed some of the goals of the project. However, such a change of the projects' objectives did not undermine the most important goal of a participatory research approach: the community members and the external researcher should share equal responsibility and power in the decision making process. Also, the project should adapt its goals and objectives to the community's needs, and not to those of the researcher.

Considering the social and economic situation of these communities, it was decided that as part of the exchange of 'knowledges' component of a participatory approach, the external researcher would facilitate English as Second Language (ESL) (or as Foreign Language, EFL) classes to the community members. This new objective was done in two different ways in the different immigrant communities that participated in the project: in the case of the community working on the promotion of Hñähñu, the external researcher divided each meeting with the community members so that half of the meeting time would be devoted to English lessons (imparted by the external researcher himself). In the case of the community members working towards the promotion of Tzotzil, the situation was different: through some university funding, the external researcher sent an L2 education student from East Carolina University to offer English lessons once or twice a week. It is worth noting (despite the small differences observed in both cases), that the English classes project had the added benefit of meeting a need of not only the language speaking community, but also of the immediate immigrant community.

Finally, in this reshaping of the research approach, all the participants agreed to elaborate the following materials: i. a basic visual vocabulary for children, with a glossary in the working language, in Spanish, and in English (that could also be of benefit to the other members of the community); and ii. a collection of traditional and immigration stories. It was agreed that all the materials produced belong to the community; that is, once completed, the materials are to be returned to the community, who will decide how to benefit from them.

In sum, the reshaping of the research approach was done always taking into consideration that the focus of such an approach has to be on the creation of a product that is beneficial to the community. Therefore, considering this focus and a somewhat lesser degree of participation by the members of the community while still maintaining a balance in the decision making, we can conclude that both projects in Eastern North Carolina better correspond to the label of a community-engaged research model.

6. Summary and conclusions

In this paper I have taken a close look at two different but related projects dealing with the promotion of immigrant indigenous languages in Eastern North Carolina and, more specifically, at the research approach applied. In both cases, the initial objective was to implement a participatory fieldwork linguistic research methodology, following the guidelines of a similar project in the promotion of Mayangna in Nicaragua, which has proven to be successful.

These projects on the promotion of Tzotzil and Hñähñu among the members of two different immigrant communities in Eastern North Carolina present evidence of two facts: first, Mesoamerican languages are part of the linguistic landscape of Hispanic immigrant populations, although their precarious sociolinguistic situation (with extremely little use of these languages outside the home and no continued use of the language in younger generations) makes evident the need to document and strengthen their language in immigrant settings; and second, and more importantly, different patterns of contact (one generated within the community, the other generated by the external researcher) can lead to similar community-engaged projects.

An early evaluation of the research methodology implemented in both projects manifests the need to distance the initial idea of a participatory research, given the fact that these immigrant communities present a set of features that are different from those observed in origin communities: i. presence of different languages and linguistic competences; ii. lower priority for language documentation efforts; and iii. different levels of literacy among community members. These differences led to a different research approach, but still maintaining the necessity to focus on the community needs and to maintain a joint decision making process. As a result, a community-engaged approach is implemented, which is based on a joint decision of the project's goals and which focuses on the adaptation of the whole project to the actual needs and priorities of the community, and on the creation of an end-result that first and foremost has to be beneficial to the community. With these priorities in mind, the team (i.e. community members and the external researcher)

reached the consensus to offer ESL classes to the members of the community, and to develop a basic children visual dictionary and a collection of traditional and immigration stories; these latter materials would serve the goal of promoting and contributing to preserve the community's indigenous language in their current immigrant setting. It is worth noting that at the time of this paper the elicitation of the data needed for the elaboration of the agreed materials was not complete; once it is completed and the materials published, they will be distributed for the community's use and benefit.

In conclusion, the two projects on the promotion of Tzotzil and Hñähñu in Eastern North Carolina make evident that working on the promotion of immigrant indigenous languages within immigrant populations following a participatory model is possible as long as the community and its needs are kept at the center of attention. Also, these two small projects manifest that, always considering the will of the members of the community, there is a real need to create similar activities and develop analogous projects to promote the visibility and survival of immigrant indigenous languages not only in Eastern North Carolina, but also wherever significant pockets of immigrant populations are settled.

References

Bach, Emmon. 1995. Endangered Languages and the Linguist. In *Linguistics in the Morning Calm III: Selected Papers from SICOL 1992*, Ik-Hwan Lee (ed.) 31–43. Seoul: Seoul Hanshin.

Benedicto, Elena & Mayangna Yulbarangyang Balna. 2007. A model of participatory action research: The Mayangna linguists' team of Nicaragua. In *Proceedings of the XI FEL Conference on 'Working Together for Endangered Languages – Research Challenges and Social Impacts'*, 29–35. Kuala Lumpur, Malaysia: SKET, University of Malaya and Foundation for Endangered Languages.

Benedicto, Elena, Dolores, Modesta & McLean, Melba. 2002. Fieldwork as a participatory research activity: The Mayangna linguistic teams. In *Proceedings of BLS 28*. Berkeley CA: University of California at Berkeley.

Burke, Garance. 2002. Yucatecos and Chiapanecos in San Francisco: Indigenous Immigrants Form Communities and enter a declining labor market. Paper presented at *Indigenous Mexican Migrants in the US: Building Bridges between Researchers and Community Leaders*. Santa Cruz, CA. October 11th–12th.

Christie, Patrick; Bradford, David; Garth, Ray; González, Bonifacio; Hosteler, Mark; Morales, Oswaldo; Rigby, Roberto; Simmons, Bertha; Tinkam, Eduardo; Vega, Gabriel; Vernooy, Ronniie & White, Noreen. 2000. *Taking Care of What we Have. Participatory Natural Resource Management on the Caribbean Coast of Nicaragua.* Managua and Ottawa: Centro de Investigación y Desarrollo de la Costa Atlántica (CIDCA) and International Development Research Centre (IDRC).

Cornwall, Andrea & Jewkes, Rachel. 1995. What is participatory research? *Social Sciences & Medicine* 41(12): 1667–1676.

Czaykowska-Higgins, Ewa. 2009. Research models, community engagement, and linguistic fieldwork: Reflections on working within Canadian indigenous communities. *Language Documentation & Conservation* 3(1): 15–50.

Eggleston, Alyson. 2012. Language and Space: A Crosslinguistic Comparison of Spatial Reference Preference in Sumu-Mayangna, Nicaraguan Spanish, and Barcelonan Spanish. Ph.D. dissertation, Purdue University.

Fink, Leon. 2003. *The Maya of Morganton. Work and Community in the Nuevo New South*. Chapel Hill NC: The University of North Carolina Press.

Fox, Jonathan. 2004. Reframing Mexican Migration as a Multi-Ethnic Process. Ms, University of California, Santa Cruz.

Hagey, Rebecca S. 1997. The use and abuse of Participatory Action Research. *Chronic Diseases in Canada* 18(1): 1–4.

Lewis, M. Paul (ed.). 2009. *Ethnologue: Languages of the World*, 16th edn. Dallas TX: SIL International. ⟨http://www.ethnologue.com/⟩

Lewis, M. Paul, Simons, Gary F. & Fennig, Charles D. (eds). 2013. *Ethnologue: Languages of the World*, 17th edn. Dallas TX: SIL International. ⟨http://www.ethnologue.com⟩

Ortega, María. 2010. The unspoken truth: Indigenous languages in the Hispanic community of Edgecombe County. Paper presented at the 7th *TALGS Conference*, East Carolina University, Greenville, NC, February 20.

Pick, Hannah, Wolfram, Walt & López, Jacqueline. 2011. Indigenous language students from Spanish-speaking countries: Educational approaches. *Heritage Briefs Collection*. Washington DC: Center for Applied Linguistics.

Rice, Keren. 2004. Ethical issues in linguistic fieldwork: An overview. *LSA*. ⟨http://www.chass.utoronto.ca/lingfieldwork/pdf/2.pdf⟩

U.S. Census Bureau. 2000. *American Fact Finder*. Washington DC.

U.S. Census Bureau. 2010. *American Fact Finder*. Washington DC.

Viñas-de-Puig, Ricard, Mayangna Yulbarangyang Balna & Benedicto, Elena. 2012. Linguistic and technical training as a community empowerment tool: The case of the Mayangna linguists' team in Nicaragua. *International Journal of Language Studies* 6(1): 77–90.

Whyte, William F., Greenwood, Davydd & Lazes, Peter. 1989. Participatory action research: Through practice to science in social research. *Behavioral Scientist* 32(5): 513–551.

Code-switching in an Erzya–Russian bilingual variety

An "endangered" transitory phase in a contact situation

Boglárka Janurik
University of Szeged, Hungary

In this paper, I study mixed code-switched structures in bilingual discourse in the Erzya-Russian code-switching (ERCS) variety on the basis of unstructured interviews. I argue that the matrix language frame model (Myers-Scotton 2002) cannot be applied to this variety, as the matrix language of the clauses cannot be identified unambiguously. I focus on verbal constructions, numeral phrases, and time expressions, which are the most typical cases of composite structures having a bilingual grammatical frame. The analysis of these constructions shows that ERCS clauses have a composite matrix language and display traces of the language mixing phase on Auer's 1999 continuum.

1. Introduction

In this paper, I examine the grammatical structure of intrasentential switches in the Erzya-Russian code-switching (ERCS) variety, the colloquial norm used by Erzya–Russian bilingual speakers in their everyday interactions. I argue that this variety shows characteristics of mixed codes. My aim is to analyze the structures that indicate the emergence of a composite matrix language, meaning that there is more than one language determining the morphosyntactic structure of the CP (complementizer phrase, the maximal projection of the complementizer) in the ERCS variety. This language mixing phase can usually be considered transitory in contact situations. The mixed structures typical of this phase provide unique data concerning the languages in contact, as code-switching "helps to account for what morphosyntactic structures are more likely to be converged and why" (Zabrodskaja 2009: 32). The informal communication of Erzya–Russian bilinguals is no longer characterized by Erzya monolingual language use. The influence of Russian is increasing, which results not only in language alternation involving the insertion

of longer stretches of Russian into Erzya discourse, but also in the emergence of phrases with double morphology and with composite Erzya–Russian grammatical structures (especially in verbal constructions and numeral phrases). Thus, studying ERCS provides important insights into composite structures resulting from the incongruence between the two participating languages in bilingual discourse. Code-switching situations typically represent clear-cut cases in which the matrix language and the embedded language can be identified unambiguously. In the case of the ERCS variety, the matrix language of the CPs is still predominantly Erzya, but there are instances in which either Russian, or Erzya and Russian together, provide the grammatical frame for the structure. In this paper, I focus on these composite structures.

First, I discuss the main characteristics of mixed codes and the processes leading to the formation of such varieties (Section 2.1). I also provide a brief overview of the categorization of code-switching types identified by the Matrix Language Frame (MLF) model (Myers-Scotton 2002) and Muysken (2000), as Muysken's 2007 classification of mixed codes is based on this. Furthermore, I introduce Auer's 1999 continuum model, which provides a subtle explanation for the evolution of mixed codes through the phases of code-switching, language mixing, and fused lects. I claim that the ERCS variety shows traces of the language mixing phase in Auer's 1999 continuum and displays structures having a bilingual grammatical frame, a composite matrix language as defined by Myers-Scotton (2001).

Second, I discuss the socio-historical background of the community (Section 2.2) and I give an overview of the typological characteristics of both Erzya and Russian (Section 2.3). The earlier sources documenting the mixed variety suggest that the ERCS variety has existed for almost a century and can be considered a relatively stable variety. However, the Erzya language, and consequently its ERCS variety, is endangered, as the language shift to Russian is the prevailing tendency due to urbanization, lack of education in Erzya, etc. (I discuss these factors in detail in Section 2.2) I also point out similar tendencies in other (Finno-Ugric) languages in contact with Russian and suggest that these phenomena should be studied in a common framework.

Third, I briefly discuss the methodology of data collection and analysis and describe the corpus in Section 3, which is followed by the detailed analysis of mixed structures of the ERCS variety in Section 4. I specifically focus on constructions in which grammatical structures of the Erzya and the Russian languages are combined. For instance, I study gender agreement in verbs, numeral phrases, and time expressions in detail. Having considered these phenomena, I argue that the ERCS variety can be considered a mixed code.

Finally, I provide an outline of further research possibilities concerning this mixed variety. I especially emphasize the role of quantitative studies, which could provide more data on sociolinguistic variation in the ERCS variety.

2. Theoretical background

2.1 The characteristics of mixed codes

The definition of mixed languages varies in publications on contact linguistics. In this paper, I am going to use Muysken's typology (Muysken 2007) and Auer's continuum model (Auer 1999).

Muysken (2007: 315) defines a mixed code as "a way of speaking which shows evidence of substantial amounts of morpho-syntactic and/or lexical material from at least two different languages". He differentiates between thirteen mixing patterns on the basis of formal distinctions. Five of these patterns refer to languages that are traditionally regarded as bilingual mixed languages (Media Lengua, Michif, Mbugu, Copper Island Aleut, and Australian mixed codes). Other subtypes of mixed codes include mixed pidgins, trade jargons, creoles, slang, and jargon-type relexified varieties, etc. From the point of view of the ERCS variety, mixed codes arising as the result of code-switching are relevant. Muysken differentiates between four main types of switching (switching is used as a broader term and mixing refers to intrasentential switching): insertional, alternational mixing, discourse marker switching, and congruent lexicalization.

I am going to define different subtypes of code-switching on the basis of the typology created by Muysken (2000). In the ERCS variety, all of the following types occur, but it is congruent lexicalization that prevails in the interviews.

Insertional switching involves the insertion of elements of language B into "a frame constituted by the rules of language A" (Muysken 2007: 320). In the Matrix Language Frame model (Myers-Scotton 2002), this inserted language is the embedded language which adds elements to the matrix language frame. In this case, it is vital that the structures of the two languages are equivalent to some extent; there should be congruence at least between the characteristics of "the inserted element and the properties of the slot into which it is inserted" (Muysken 2007: 320).

These inserted elements are integrated into the matrix language by matrix language elements – and usually resemble borrowings – as they are either one-word constructions or fixed phrases. In the case of the Erzya–Russian bilingual community variety, the differentiation between borrowings and code-switches is quite impossible, as the usual criteria (phonological and morphological adaptation, frequency of usage, acceptance by the community) do not yield a clear-cut distinction between these categories. The similar phonological systems of the two languages, the lack of monolingual speakers, and insufficient data on Erzya–Russian contact phenomena, do not enable an unambiguous differentiation between borrowing and code-switching. In my analysis, I am going to take all Russian elements occurring in the corpus into consideration and mark them in bold in my examples below.

In Muysken's typology, the second type of intrasentential code-switching is alternational mixing. In this case, one clause is taken from language A, while the second clause is from language B. This type of switching is especially common in cases where the grammatical structures of the two languages involved do not fit together, and elements from these languages are thus juxtaposed rather than combined. Clause-peripheral elements are common in alternational switching. Adverbial phrases, for instance, are frequently switched this way, and the switching usually has a foregrounding function (e.g. time expression in my corpus, cf. Example 13). The same tendency can be detected even in the case of discourse markers. (This latter type of mixing is sometimes considered to be a separate subtype of switching.)

The last subcategory of intrasentential switching in Muysken's classification is called congruent lexicalization, which is "the rapid back and forth switching of loose elements in a structure mostly shared by the two languages" (Muysken 2007: 322). According to Muysken, this type of contact phenomenon is rather common in cases where the switching occurs between standard and dialect varieties, closely related languages, or if there is significant convergence between the languages. The ERCS variety represents the third of these situations. According to Aikhenvald (2006: 47), "[i]n the situation of one language dominating the others, convergence may involve gradual adoption of the other language's structures at the expense of its own". The grammatical system of Erzya has become more similar to the grammatical system of the Russian language throughout the centuries of language contact that enabled the emergence of congruent lexicalization in contemporary Erzya. (I will give an overview of the contact situation in Section 2.2)

In Muysken's theory, distinct code-switching types are typical of given contact situations. If social circumstances change, contact phenomena, their usage, and function also tend to change. Auer represents this process in a continuum model which involves three phases: code-switching → language mixing → fused lects. This transition can be considered as a type of grammaticalization process. In the first phase, code-switching is optional and when it is applied, it always has a distinct pragmatic function. This starting stage can be described by the MLF model and is called classic code-switching. In the language mixing phase, we find variation where parallel forms co-exist, and the matrix language of the discourse cannot be defined unambiguously. In the MLF model, the term *composite matrix* is applied to this bilingual grammar, as clauses are constructed relying on the rules of both languages. There are, however, negative grammatical constraints that block mixing at certain points (Sarhimaa 1999: 148). These restrictions are usually variety-specific. For instance, in the ERCS variety it is possible to switch between the subject and the predicate, but no switching occurs in negative constructions between the negative particle and the verb.

The code-switching phase involves both insertion and alternation. Code-switching has pragmatic functions: dominant language translations in the form of insertions are applied by speakers to ensure that meaning is conveyed; alternational code-switching is typically used for citation (in reported speech), emphatic repetitions, and summaries, etc. (Cf. Auer (1999: 321) for an extensive list of these functions.)

Due to frequent use of code-switching forms (both insertions and alternations) the discourse structuring function of the embedded language elements might be lost. The increased inflow of L2 elements to L1 discourse makes the status of the matrix language ambiguous. The MLF model (Myers-Scotton 2002) also explains the process of how the more frequent application of the embedded language can lead to change in the matrix language. If the number of insertions grows – especially if it involves elements with embedded language grammatical markers – it might lead to an increase in the influence of the embedded language. In Myers-Scotton's 2002 model, the change in the matrix language status, i.e. when the former embedded language becomes the matrix language of the discourse, is called Matrix Language Turnover. In social circumstances that facilitate language maintenance, the mixed variety fossilizes and becomes the community's main variety with a composite matrix language (CML):

> The notion of a CML is a broad one, characterized mainly by 'long embedded island switches', and what we term semantic, syntactic and morphological transference, which are disparate phenomena and may or may not be interrelated in particular instances. (Clyne 2003: 136)

The last stage of the process is that of the fused lects level, where there is no variation: the use of a form from the former L1 or L2 is obligatory for a given grammatical function. The code-switching → language mixing → fused lects transition might stop at any phase and the mixed variety can become the community language.

Speakers of the ERCS variety produce constructions that are characteristic of the language mixing phase. Variation is also present in the ERCS variety, especially in verbal constructions. Numeral phrases, especially involving compound numbers, however, are used exclusively in Russian, which is rather typical of the fused lects phase.

2.2 Sociohistorical background

As mixed codes predominantly emerge under certain social conditions, it is crucial to characterize the sociohistorical background of the Erzya–Russian bilingual community in order to understand the linguistic picture.

Erzya belongs to the Mordvin subgroup of Finno-Ugric languages. Its status has been intensely debated: there is no agreement whether it should be considered

a Mordvin dialect (along with Moksha), or a separate language. There is no reliable data even on the number of Erzya speakers. According to the 2010 Russian census (Russian census 2010), there are 744,237 ethnically Mordvin (Moksha and Erzya) people in the Russian Federation, of whom 392,941 are speakers of one of the Mordvin languages. Informal estimates suggest that Erzya speakers constitute about two-thirds of all Mordvin speakers. Only 33 percent of the Mordvin speakers live in the Mordvin Republic (Keresztes 2011: 11); most of the speech community is in a diaspora situation in which the number of language domains where the mother tongue could be used is limited. In the Mordvin Republic, the Mordvin languages are *de jure* official languages. However, this does not mean that they *de facto* have an equal status with Russian. The purist attitudes of the Mordvin intelligentsia also contribute to the tendency of language shift. As a result of the stigmatization of the ERCS variety, speakers incapable of producing monolingual Erzya speech shift to Russian and do not transmit the language to the next generations. If we compare data from the 2002 census (Russian census 2002) and the 2010 census (Russian census 2010), we can see a radical decline in the number of Mordvin speakers: 614,260 in 2002 and 392,941 in 2010.

In this paper, I regard Erzya and Moksha to be separate languages, taking into consideration their diverging development since their separation in the 13th century. Furthermore, two standard languages (standard Moksha and standard Erzya) were established for the two varieties in the 1920s due to the Soviet language policy and separatist intentions of the Moksha and Erzya intelligentsia.

Among the Finno-Ugric languages, it was the Mordvin languages that first came into contact with Russian varieties. Early trade connections were established in the 9th century as Russian settlers arrived in the Mordvin areas, with contacts intensifying in the 11th century (Keresztes 2011: 15). The first two centuries of the contact situation were characterized by very few loanwords, as Mordvin groups under the most prevailing Russian influence assimilated into the Russian community. However, this long-term contact situation has since resulted in extensive borrowing and structural changes in the Mordvin languages.

The current dispersion of the Mordvin communities can be traced back to Russian administrative and religious policies in the late Middle Ages. In the 16–17th centuries, the Mordvin groups moved eastwards to the Transvolga region and beyond, escaping high taxes and the spread of the Orthodox Christian faith (Keresztes 2011: 15). However, the entire Mordvin community was Christianized by the 18th century.

Diaspora communities living in the Ulyanovsk, Samara, and the Orenburg Oblasts 'provinces', the Chuvash and Tatar Republics, and also in the Ural Mountains have more or less close connections to the Mordvin Republic. New ways of cooperation have been created especially after the fall of the Soviet Union in 1991. Teacher conferences and joint celebrations of Erzya and Moksha communities

along the Volga River serve unifying purposes. Among the Mordvin intelligentsia, there are intentions to create a unified Mordvin standard variety in order to increase the chances of first language maintenance in all Mordvin speakers.

It was the diaspora speakers and members of the intelligentsia who first became bilingual in one of the Mordvin languages and Russian at the end of the 19th century. Russian contact phenomena were detectable both in their written and spoken varieties. To some extent, these structures are similar to the code-switching constructions in the contemporary ERCS variety.

Characteristic features of the ERCS variety can also be detected in the case of other minority languages in contact with Russian. There have only been a few studies carried out on these mixed varieties. Extensive research is available only on Lovozero Sami (Pineda 2009), Karelian (Pyöli 1996; Sarhimaa 1999), and Tatar (Wertheim 2003). As similar tendencies are detectable in these varieties – numeral phrases, time expressions, (gender agreement in) Russian finite verbal forms, etc. – further studies should focus on creating a theoretical framework for the analysis of these contact phenomena.

2.3 Typological characteristics of Erzya and Russian

In this section, I focus on typological features relevant for my study, predominantly on the typological characteristics of numeral phrases and verbal constructions in Erzya and Russian. The typological overviews are based on Raun (1988) and Zaicz (1998) for Erzya and on Timberlake (1993) for Russian.

In Russian, numeral phrases are constructed on the basis of the following rules: the nominative singular is used after the number 'one' and compound numbers ending in one (with the exception of compound numbers ending in eleven); the singular genitive is applied after the numbers 'two', 'three' and 'four', and their compound counterparts (with the exception of numbers ending in twelve, thirteen and fourteen); and the plural genitive is the required form in all other cases. The Erzya system, however, is less complex: the argument of the numeral phrase is in the nominative case; if the attribute is the number 'one', the singular form is used, whereas after numbers 'two' and bigger, a plural ending is usually required.

For the discussion of numeral phrases and especially time expressions, it is essential to give an overview of the nominal structures in both languages. In Erzya (Raun 1988: 100), nouns have three declensions: indefinite (*kudo* 'house'), definite (*kudoś* 'the house'), and possessive (*kudom* 'my house', *kudot* 'your house', etc). There are 11 cases: nominative, genitive, allative-dative, ablative, inessive, elative, illative, prolative, translative, comparative, and caritive. Apart from these case suffixes, Erzya has many postpositions, but no prepositions. Like all the other Finno-Ugric languages, Erzya has no gender distinction.

In Russian, "nominal parts of speech express distinctions of case, number and gender" (Timberlake 1993: 836). There are two numbers (singular and plural) and six primary cases (nominative, accusative, genitive, locative, dative, and instrumental); the two secondary cases – secondary genitive, and secondary locative – are used only with a small number of masculine nouns. Russian distinguishes between three grammatical genders: masculine, feminine, and neutral. (Syntactic) gender is "expressed through agreement in other parts of speech – attributive adjectives, predicative adjectives, the past tense of verbs and ultimately pronouns" (Timberlake 1993: 836).

There are two conjugations in Erzya: the indefinite and the definite. The definite conjugation is used if the verb is transitive, the direct object is definite, and the aspect of the action is perfective. Erzya distinguishes between four tenses: present, two past tenses, and a periphrastic future. (The second past refers to habitually recurring or long-lasting events in the past.)

Russian differentiates between imperfectives and perfectives. According to Timberlake (1993: 849), "[i]mperfectives distinguish past, present and future… Perfectives distinguish past and a morphological present, which reports true future or singularized habitual situations". The present is inflected for person and number, while the past is inflected for gender and number. Past tense forms and gender agreement in the past tense are discussed in detail in Section 4.1.

There are significant differences in the order of constituents in Erzya and Russian. In noun phrases, the modifier precedes the modified in both languages: Russian *starij dom* 'old house' (*starij* 'old' and *dom* 'house') and Erzya *tašto kudo* 'old house' (*tašto* 'old' and *kudo* 'house'); the same order is common also in possessive structures: *babań kudo* 'the grandmother's house' (*baba* 'grandmother' with the genitive singular suffix *-ń*). In Russian, however, the possessor follows the possessed: *dom babuški* 'the house of the grandmother' (*dom* 'house' and *babuška* 'grandmother' (in genitive singular)). In numeral phrases, the order of constituents is the same in both languages: the numeral precedes the modified. In Russian numeral phrases, however, approximation is expressed by a reversed order – the numeral follows the complement: *dva časa* 'two hours', but *časa dva* 'approximately two hours' (*dva* 'two' and *časa* 'hour' (in genitive singular)). In Erzya, approximation can be expressed by the comparative case suffix: *-ška*: *kavto čast* 'two hours' and *kavtoška čast* 'approximately two hours' (*kavto* 'two' and *čas* 'hour').

Word order in Erzya is SVO, which evolved under the influence of Russian from an earlier SOV word order. According to Zaicz (1998: 206), word order in Erzya varies: "The basic rule is: topic(s) to the head of the sentence, focalized element immediately before the finite verb; a focalized verb must therefore stand sentence-initial".

Russian word order varies, but SVO is considered dominant (WALS 2013). According to Timberlake (1993: 858), "the word order of the predicate and its

major noun phrases (subject and objects) is relatively free in Russian". The different word orders can have different stylistic functions and also: "[t]he naturalness and frequency of various orders depends on the role of the noun phrase and the semantics of the verb" (Timberlake 1993: 858).

3. Methodology and the corpus

In this paper, I discuss data collected during field trips to the Mordvin and Chuvash Republics in 2008, 2010, and 2011. I studied Erzya communities in two regions of the Mordvin Republic and in one diaspora area in the Chuvash Republic. Locations in the Mordvin Republic were selected on the basis of the status of the Erzya community: its majority or minority status in the village/town in question. Baevo is an Erzya-majority village (1906 people) with a kindergarten and school providing Erzya-language education (which is only provided in villages with an Erzya majority), and Aťashevo is a small Russian-majority town (5900 people). Two villages (Atrať with 1026 people and Altishevo with 1548 people) were chosen from the diaspora region, i.e. from outside the border of the Mordvin Republic. These villages represent a unique situation among the Erzya communities as their speakers are doubly minorities in the places where the dominant languages are Russian and Chuvash. However, these diaspora communities located in areas bordering the Mordvin Republic are still in a better sociolinguistic situation than diaspora speakers in more remote regions like the Ulyanovsk community living in the Ulyanovsk province, as Erzya-language publications and radio and television programs are available to them.

The corpus contains structured and unstructured interviews and spontaneous conversations from a total of 20 speakers: five consultants were selected from each of the four locations. The number of male speakers is rather low in the sample, only three. Consequently, I am not able to study gender differences in the language use of ERCS speakers, and will not attempt to do so. I examine age, however: the majority of the consultants (14 speakers) belong to the middle-aged age group (30–60 years), while the remaining six speakers are representatives of the elder generation (over 60 years). Only two of the speakers have university or college degrees, whereas most of my consultants have secondary education (14 speakers) and four only attended a primary school for a few years. The sample of consultants is not representative. I had to rely on the friend-of-a-friend method. Consequently, I do not carry out a quantitative analysis of the data in order to analyze the correlations between the sociolinguistic background of the speakers and types of contact phenomena occurring in their ERCS variety.

4. Discussion

I now want to turn to the main focus of this paper and discuss some of the most frequent mixing phenomena occurring in the ERCS variety. The discussed inventory of constructions is far from complete since my purpose is to identify the code-switching tendencies in the ERCS variety and create a solid basis for a future quantitative analysis of the data. Having discussed characteristics of mixed codes, I enlist contact phenomena in the ERCS variety that seem to indicate that ERCS can also be considered a mixed variety. Insertion, alternation, and congruent lexicalization all occur in the ERCS variety. In my corpus, there seems to be a clear difference in the frequency of these code-switching types, which varies along several factors. Still, on the basis of the data in my corpus, I can safely claim that congruent lexicalization appears to prevail in the ERCS variety. To support this claim, I outline the general code-switching tendencies in the ERCS variety, focusing on constructions formed on the basis of Russian grammatical rules. I analyze verbal constructions, numeral phrases, and time expressions in detail.

4.1 Verbal constructions

In the ERCS variety, Russian verbal forms either occur with Erzya suffixes or with Russian system morphemes. There are no examples in my corpus in which an Erzya stem would have a Russian suffix. This indicates that the Erzya language is the dominant language of the bilingual discourse. As Bentahila and Eirlys (1998: 26) put it, "Grammatical morphemes from the dominant language can be attached to lexical morphemes of the non-dominant one, whereas the reverse never happens".

The use of Russian endings does not depend on the person marking of the grammatical subject, or whether the word denoting the subject is in Russian or Erzya. In Russian, verbs are marked for person and number in the present tense whereas in the past tense, for gender or number (*-l*, *-la*, *-lo* – masculine, feminine and neutral singular; *-li* – plural) – cf. Section 2.3. ERCS sentences with a Russian verb are always marked for person in the present tense, whereas in the past tense the use of gender agreement rules varies.

In my corpus, gender agreement markers were attached to verbs only in sentences with an animate subject, which could be expressed by either a personal pronoun or a noun phrase. In the case of inanimate subjects (cf. Example 1), no gender agreement occurred in my data. Since Erzya has no grammatical gender distinction (cf. Section 2.3), the emergence of this partial gender agreement system is a sign of the development of a mixed code and that Erzya (or at least the ERCS variety) is gradually converging with Russian.

(1) ***nacionaľnost'***-*eś vadŕa* ***i*** *mońeń* ***pomog***[1]
nationality-DEF good and I.DAT.SG help.PST
'My nationality is good and it helped me out.'

In Example (1), the subject *nacionaľnost'* 'nationality' is inanimate. In Standard Russian, *nacionaľnost'* is a feminine noun which would trigger the use of the feminine suffix in the past tense form of the verb *pomoč'* 'to help', so the form *pomogla* would occur in a monolingual Russian sentence (the form *pomog* is an exception, some verbs have an unmarked masculine form in the past tense, i.e. without the past masculine ending *-l*). In the second clause (*i mońeń pomog* 'and it helped me out'), the matrix language of the clause cannot be determined as the verb obeys the rules of the Russian language, while the personal pronoun has an Erzya dative suffix. (The argument of the verb 'to help' is dative in both Erzya and Russian.)

In the following examples, I give typical examples of clauses with animate subjects.

(2) *toso meźejak araś mon* ***objasńi-l-a***
there nothing be.NEG 1SG explain-PST-F
'There is nothing, I explained.'

Example (2) illustrates a case in which the subject is a personal pronoun. It is a tendency in my corpus that subjects expressed by a personal pronoun trigger gender agreement. There was only one exception to the rule: an 86-year-old consultant who grew up monolingual and illiterate, and started to learn Russian when she was evacuated from Mordovia to Mongolia during World War II.

(3) *mon* ***ńe hoťe-l,*** ***ńe hoťe-l*** *venča-ms*
I not want-PST.M not want- PST.M marry-INF
'I did not want to get married.'

In the case of Standard Russian and in the speech of younger ERCS speakers, the verb would have the feminine past tense ending *-la* (*hoťela* 'wanted'). This constant use of the masculine form of the verb in this consultant's speech suggests that gender agreement has started to spread in ERCS only recently, thus we can observe

1. In the examples, bold face is used for highlighting the Russian elements in the clauses. Throughout this paper, the following abbreviations are used in the glosses: 1PL 'first plural', 1SG 'first singular', 3PL 'third plural', 3SG 'third singular', ABL 'ablative', ACC 'accusative', DAT 'dative', DEF 'definite', ELAT 'elative', EMP 'emphatic', F 'feminine', GEN 'genitive', INESS 'inessive', INF 'infinitive', INS 'instrumental', LAT 'lative', M 'masculine', NEG 'negation', PL 'plural', POSS 'possessive', PRS 'present', PST 'past', 1PL<3PL 'definite conjugation with the subject in third person plural and the object in first person plural', TRANSL 'translative'.

an ongoing change in the speech of Erzya–Russian bilinguals. The fact that use of gender agreement also varies in the speech of younger speakers, especially in cases involving a subject expressed by a noun, also supports this claim.

(4) *jalga-m* ***tože*** ***reši-l-a*** ***postup-ať***
friend-1SG.POSS also decide-PST-F apply-INF
'My friend also decided to apply to the university.'

In Example (4), the subject is an Erzya noun (*jalgam* 'my friend') and the verb has a feminine past tense suffix. If we compare this example to Example (8), we can see that it is not relevant if the subject has an Erzya or a Russian stem. In Example (4), the matrix language is Russian and there is only one Erzya system morpheme: the first person singular possessive suffix *-m* in the word *jalgam* 'my friend'; other system morphemes as well as the grammatical structure of the sentence are Russian.

Gender agreement is also common in auxiliaries. Russian auxiliaries (*hoťeť* 'want', *moč'* 'can') are commonly used in the ERCS variety (cf. Example 3).

(5) *mon* ***uže*** *viška ping-ste* ***hoťe-l-a*** *uľ-ems* ***učiťeľ****-eks*
1SG already small age-ELAT want-PST-F be-INF teacher-TRANSL
'I have wanted to be a teacher since I was small.'

In Example (5), the matrix language of the clause cannot be determined unambiguously. This is a typical case of congruent lexicalization (Muysken 2000) or of a composite matrix language (Myers-Scotton 2002) where the structure is based on a bilingual grammatical structure. The predicate *hoťela* 'wanted' has a Russian system morpheme (the past tense feminine suffix *-la*), but other system morphemes are in Erzya: the suffix *-ems* and the translative case suffix *-eks*. On the basis of the use of gender agreement, Russian could be defined as the matrix language of the clause. According to the MLF model, however, mixed constituents such as *učiťeľeks* (Russian stem + Erzya case suffix) should be inserted into the clause with a matrix language system morpheme, which suggests that the matrix language is Erzya.

In addition to auxiliaries such as *hoťeť* 'want', which behave as main verbs and can display gender agreement only in the past tense, there are other auxiliaries (e.g. *dolžen* 'must, have to') that can agree with the subject in gender in the present tense as well. The auxiliary *dolžen* 'must, have to', expressing necessity or obligation, has four forms in the present tense: three forms in the singular, *dolžen* (masculine), *dolžna* (feminine), *dolžno* (neutral), as well as *dolžni* in the plural.

(6) *śe kolko-ś* ***dolžen*** *rama-ms* ***vśe*** ***čemodan****-ś-gak*
that godfather-DEF have.to.M buy-INF whole suitcase-DEF-EMP
'That godfather has to buy the whole suitcase.'

In Example (6), the subject is of masculine gender and the morphologically unmarked form of the auxiliary is used. (This coincides with the masculine form of the auxiliary.) There are no instances in my corpus in which the feminine form of the verb would be used in the case of a masculine subject.

The auxiliary *dolžen* 'must, have to' can also occur in its feminine form in the ERCS sentences.

(7) *mejľe* ***ńevesta*** ***dolžn-a*** ***podarok*** *maks-oms* *ťe* *ćorakaj-śťe*
then bride have.to-F present give-INF that boy-DAT.DEF
'Then the bride has to give presents to that boy.'

The use of the Russian auxiliary and the use of the feminine form could be triggered by the fact that the subject is also expressed by a Russian word. However, there are other instances (cf. Example 8) where the Russian subject form does not trigger gender agreement.

(8) ***učiťeľńica****-nok* ***dolžen*** *sa-ms*
teacher.F-1PL.POSS have.to.M come-INF
'Our teacher is supposed to come.'

In Example (8), there is no gender agreement; the auxiliary is in the unmarked form despite the fact that the subject is the feminine form of the word 'teacher'. The clause has a bilingual grammatical structure, as it contains system morphemes from both languages: the possessive suffix (*-nok*) and the infinitive ending (*-ms*) from Erzya and the morphologically unmarked form of the auxiliary from Russian. The mixed constituent *učiťeľńicanok* 'our teacher' is inserted into the clause using an Erzya possessive ending, which would indicate that the matrix language is Erzya. However, even though there is no gender agreement in the clause, the fact that the predicate is formed in accordance with Russian rules suggests that Russian is the matrix language of the clause.

Gender agreement is also sometimes marked on Russian adjectives in ERCS.

(9) *ťeťa-ń* *jondo* *baba-m* *uľńe-ś* *son*
dad-GEN.SG side grandmother-1SG.POSS be-3SG.PST 3SG
uľńe-ś *pek* ***strog-aja***
be-3SG.PST very strict-F
'My grandmother on my father's side was very strict.'

In Example (9), the Erzya sentential subject *babam* 'my grandmother' and the anaphoric pronoun *son* '(s)he' is linked with a predicative adjective *strogaja* 'strict', which is marked for feminine gender. The structure again follows the agreement rules of the Russian language, although the matrix language of the sentence is clearly Erzya.

In this subsection, we could conclude that the use of gender agreement does not depend on the matrix language of the clause. Animacy and the form of the subject seem to determine whether gender agreement occurs in the clause. The variation in the use of gender agreement (and especially its absence from the speech of the older generation) suggests that its emergence represents an ongoing change in the ERCS variety.

4.2 Numeral phrases

In addition to verbal constructions, numeral phrases (especially in time expressions) and prepositional phrases are typically code-switched in the ERCS variety. The insertion of Russian numeral phrases into Erzya discourse has been a long-standing practice among Erzya-Russian bilingual speakers. This tendency could already be observed in the 1960s, cf. Hallap's article: "expressions referring to dates, time, quantity, length, weight, etc. are usually expressed using only Russian structures" (Hallap 1960: 222; my translation). Bearing in mind the evidence in Sarhimaa (1999) and Pineda (2009), we can see that the use of Russian numerals (especially compound numbers) is not an isolated phenomenon in the ERCS variety; it is also found in other minority languages in contact with Russian.

Most of the numeral phrase types occurring in the corpus can be considered embedded language islands (according to the MLF model) where the head (the numeral) has a nominal complement in a form required by the rules of the Russian language. For example, *šem'deśat vośem' let* ('seventy-eight years'; *let* is the suppletive genitive plural form of the word *god* 'year'). In some other cases, however, the complement fails to follow the Russian rules and occurs in the nominative case where the usage of this case would be erroneous in Russian: *sorok četiŕe god* ('forty-four years', lit. 'fourty-four year'). Mixed structures occur when the Russian number is followed by the Erzya equivalent of the word, or when the head is an Erzya number and is accompanied by a Russian complement.

The following example represents some of the characteristic types of code-switches that I will discuss in detail in Sections 4.2.1 and 4.2.2 below:

(10) *ŕad-i-ńek* ***šesť*** ***meśac-ev*** *toso* ***učebnoj otŕad****-so,*
finish-PST-1PL six month-GEN.PL there training battalion-INESS
tonavt-imiź ***na*** ***specialʹnosť*** ***torpedist-a***
teach-PST.DEF.1PL<3PL to specialty torpedo.operator-GEN.SG
i *mejľe* ***značit*** *pong-i-ń* ***Kamčatka-v*** ***na podvodn-uju***
and then well get-PST-1SG Kamchatka-LAT to
lodk-u ***na*** ***raketn-uju*** ***i*** ***vot*** *toso*
underwater-ACC ship-ACC to rocket-ACC and so there

služ-i-ń *kolmo* ***god-t.***
serve-PST-1SG three year-PL

'We finished the training camp in six months, we were trained as torpedo operators, and then, well, I was assigned to Kamchatka to a submarine (with rockets) and so I served there for three years.'

In Example (10), the two time expressions *šesť meśacev* 'six months' and *kolmo godt* 'three years' represent two distinct types of constituents. The structure *šesť meśacev* 'six months' is a Russian embedded language island, as it is formed according to the rules of Russian and it is inserted into an Erzya clause. The construction *kolmo godt* 'three years' is a mixed constituent in which the Russian word *god* 'year' receives the Erzya plural ending *-t*, as is required in structures having numbers 'two' and bigger as attributes.

In the following two subsections (4.2.1 and 4.2.2), I give an overview of cases in which numeral phrases in the ERCS variety are formed following standard Russian rules or in which they show traces of a mixed-language grammar.

4.2.1 *Quantity*

Typological characteristics of the Russian and Erzya numeral phrases were presented in Section 2.3. In ERCS structures, these two systems are combined. Accounts of the ERCS variety draw the borderline of the native and the Russian numeral systems at different points. For instance, Agafonova (2002) claims that diaspora speakers use Erzya numerals up to '10'. In my corpus, the borderline is usually around '5' or '6'.

With the following examples, I intend to show how the Erzya and the Russian numeral systems are combined.

(11) *śe* ***plaťija-ś*** *kis* *mońe* *maks-śť* ***četiŕe***
this dress-DEF for me.DAT.SG give-3PL.PST four

kilogramm-a *tovźuro-t*
kilogram-GEN.SG grain-PL

'They gave me four kilograms of grain in return for the dress.'

In Example (11), the 'four kilograms' part of the 'four kilograms of grain' structure obeys the rules of the Russian language, while the Erzya word 'grain' is in plural in concordance with the Erzya system.

Examples (10) and (11) can be analyzed by the MLF model, as the matrix language in both cases is unambiguously Erzya. Embedded language elements can occur only as embedded language islands or in mixed constituents which are constructed obeying the grammatical rules of Erzya, the matrix language. This agrees with the Myers-Scotton's 2002 model of classical code-switching.

However, there are mixed structures in my corpus which have a bilingual grammatical structure. In these phrases, the Erzya head has a Russian complement and its form follows the rules of the Russian language (cf. Example 12):

(12) *mińek viŕ-eńek* ***naverno kilometr-a***
our forest-1PL.POSS perhaps kilometer-GEN.SG
kavto ej-ste-ďe-ńek
two us-ELAT-ABL-1PL.POSS
'Our forest is perhaps two kilometers from us.'

The phrase *kilometra kavto* 'around two kilometers' combines the rules of the two languages. In Russian, approximate quantities are expressed by an inverse order of the head and the complement of the numeral phrase; i.e. the complement precedes the head in these instances, so the Russian form in monolingual speech would be *kilometra dva* 'around two kilometers'. In Erzya, however, approximation is expressed by the suffix *-ška*. In standard Erzya, 'around two kilometers' would be expressed as *kavtoška kilometrat* (with the complement in plural). The word *kilometra* can either be in the nominative, as in Erzya, or in the Russian genitive singular form (of the word *kilometr* 'kilometer'). According to Russian rules, the number 'two' requires its complement to be in the genitive singular. In any case, the phrase in Example (12) must be analyzed as a mixed structure as it either adapts to the rules of both languages or to a composite matrix language.

4.2.2 *Time expressions*

Time expressions are mentioned among the most common instances of code-switching in all accounts of the ERCS variety, and also typically of other bilingual mixed varieties. Pineda detects the same tendency in Lovozero Sami: "The year of birth and age are almost always given in Russian" (Pineda 2009: 33; my translation).

Time expressions can involve adverbials of time (e.g. *rańše* 'earlier'), prepositional phrases (e.g. *do sih por* 'to this day'), noun phrases (e.g. *pervij raz* 'for the first time'), and affixed constructions (e.g. *v ďevjanosto vtorom godu* 'in 92').

In the case of adverbials of time, the mixed grammatical system of the two languages is not as detectable as in the case of prepositional phrases or affixal constructions, as these adverbials (*śejčas* 'now, at the moment' in Example 13) tend to occupy a marginal position at the beginning of a clause, and thus are instances of alternational switching rather than representatives of the insertional type of code-switching.

(13) ***śejčas*** *mon eŕa-n* ***pośolok-so***
now 1SG live-1SG.PRS small.town-INESS
'I live in a small town at the moment.'

Time expressions in affixal and prepositional constructions are either embedded Russian language islands in the Erzya clause, or Erzya suffixes are attached to Russian stems, forming mixed constructions. Example (14) represents a typical case of how ERCS speakers express dates.

(14) *ťeze* ***perejeha-l-i*** ***v šesťdešat*** ***šeďm-om*** ***god-u*** *ťese* *eŕa-tanok*
to.here move-PST-PL in sixty seven-INS year-INS here live-1PL.PRS
uš ***sorok*** ***tri*** ***god-a***
already forty three year-GEN.SG
'We moved here in '67 and we have already lived here for forty-three years.'

These two time expressions (*v šesťdešat šeďmom godu* 'in '67' and *sorok tri goda* 'forty-three years') both obey the rules of the Russian language. In the first clause, Russian is the matrix language. In the second, however, Erzya provides the grammatical frame for the clause, and consequently is the matrix language. The time expression *sorok tri goda* 'forty-three years' is an embedded language island. However, not all the cases of code-switched Russian time expressions are this unambiguous.

In mixed constructions, Erzya suffixes are attached to Russian structures. In Example (15), the Erzya inessive suffix *-se* appears on a construction formed on the basis of Russian grammatical rules (after the number '18', the suppletive genitive plural form *ľet* of the word *god* 'year' is used in Russian):

(15) *miŕďe-ńeń* *ľiś-i-ń* ***vośemnadcať*** ***ľet-se***
husband-DAT go-PST-1SG eighteen year-INESS
'I got married when I was eighteen.'

Hallap (1960) mentions that this mixing tendency can have more extreme forms in cases where a Russian ending is attached to Erzya/Moksha words. I have no examples of this tendency in my corpus. However, in a smaller database collected from diaspora speakers from the Ulyanovsk region, there are examples of this phenomenon (cf. Example 17).

(16) ***da-jut*** *lamo* ***ṕribiľ'*** ***dľa…*** *veľe-ńťeń*
give-3PL.PRS many profit to village-DEF.DAT
'They give a lot of profits… to the village.'

Example (16) represents a case of double marking where the Russian preposition (*dľa* 'to/for') and the Erzya suffix (*-ńťeń* 'to/for the', the semantic counterpart of the Russian preposition) also occur with an Erzya stem. This is, however, not a smooth switch: the speaker pauses after the preposition indicating the discrepancy between the structures in the two languages.

The other case, given in Example (17), is a clearer instance of the matrix language turnover, with the noun *veľe* 'village' receiving the Russian plural genitive ending:

(17) *te veľe-se-ńť mińek* ***značit*** *uľ-i* ***ozero*** *kov*
this village-INESS-DEF our so be-3SG.PRS lake to.where
loma-ť jak-iť ***otdih-ať i*** ***prijezža-jut*** ***iz***
person-PL go-3PL.PRS rest-INF and come-3PL.PRS from
drug-ih *veľe-j*
other-GEN.PL village-GEN.PL
'So, in this village we have a lake where people go to rest, they come from other villages.'

This phenomenon (attaching Russian endings to Erzya stems) is the opposite of cases presented above: mixed constructions in this diaspora discourse (e.g. the *iz drugih velej* 'from other villages' in Example 17) are inserted into the clause by using Russian morphological elements. This feature is only one indicator of the ongoing language shift in the diaspora communities, accompanied by the use of bare forms and the attrition of the grammatical system which are clear signs of the matrix language turnover.

The double marking strategy in Example (16) is replaced by the exclusive use of Russian system morphemes in Example (17). Russian is the matrix language into which Erzya elements are inserted. Although the first clause in Example (17) (*kov lomať jakiť otdihať* 'where people go to rest') has Erzya as its matrix language, and could represent a typical ERCS clause, the structure of the second clause indicates that the matrix language turnover has already occurred or it is an ongoing change in this diaspora variety. In my corpus of ERCS discourse, there are no cases in which mixed structures would have Russian morphology. Russian system morphemes are applied only to Russian stems. It is only Erzya system morphemes that can be attached to both Erzya and Russian stems. Although the influence of Russian on Erzya is increasing and Erzya is more and more converging with Russian, applying the terminology used by Bentahila and Eirlys (1998: 26) (cf. Section 4.1), we can claim that Erzya is still the dominant language in the ERCS variety.

5. Conclusions

On the basis of the mixed constructions discussed in detail above, I consider the ERCS variety a bilingual mixed variety which shows characteristics of the language mixing phase on Auer's 1999 code-switching → language mixing

→ fused lects continuum. Conversely, earlier studies on Russian contact phenomena in Erzya (and Moksha) suggested that the level of mixture is typical of the code-switching phase on this continuum. Additionally, diaspora data from my database and also examples presented by Agafonova (2002) imply that, for sociolinguistic reasons (specifically, the few domains where the community language is used and the absence of school education in Erzya, etc.), the Transvolga varieties started to show traces of language shift (and a turnover of the matrix language).

The mixed constructions presented above, however, do not indicate that matrix language turnover would have occurred in the ERCS variety, as Russian system morphemes are predominantly applied to Russian stems. Erzya is clearly the dominant language in the ERCS discourse; however, the matrix language of the clauses in my corpus could not be defined unambiguously in all cases. Examples for congruent lexicalization in which the clause is based on a bilingual grammatical frame were found in verbal constructions with gender agreement (cf. Example 5), numeral phrases (cf. Example 12), and time expressions (cf. Example 15).

If the social and political conditions in the Mordvin Republic favor the maintenance of the Erzya language in the future, it might result in the elimination of Russian forms even from the spoken variety of the language, leading to a minimal code-switching variety of Erzya, or the mixed ERCS variety might fossilize as the community's main language. Although the ERCS variety has existed for over a century, its status changes constantly. The documentation of this transitory phase is vital for a better understanding of language change and language contact.

As the next step of my investigations into ERCS, I intend to carry out a more extensive and more detailed analysis of mixed structures in the speech of Erzya–Russian bilinguals. I believe that the study of the ongoing change leading towards a composite matrix and a mixed language mode will enhance our knowledge of language contact and languages in general, and, most importantly, the structure of the endangered Erzya language.

In order to provide a more accurate account of this mixed numeral system, a questionnaire survey should be conducted, which would test to what extent Erzya speakers can produce Erzya numerals, and to what extent they feel comfortable using them.

I also intend to conduct detailed quantitative analyses of the ERCS variety that might reveal how the negative grammatical constraints that block mixing at certain points (cf. Sarhimaa 1999: 148) function in this variety.

References

Agafonova, Nina. 2002. Роль языковых контактов в возникновении инноваций в смешанных мордовских говорах Заволжья (The role of language contacts in the emergence of innovations in the mixed dialects of the Transvolga region). In *Volgan alueen kielikontaktit*, Jorma Luutonen (ed.), 65–69. Turku: University of Turku.

Aikhenvald, Alexandra Y. 2006. Grammars in contact: A cross-linguistic perspective. In *Grammars in Contact: A Cross-linguistic Typology*, Alexandra Y. Aikhenvald & Robert M.W. Dixon (eds.), 1–66. Oxford: OUP.

Auer, Peter. 1999. From codeswitching via language mixing to fused lects: Toward a dynamic typology of bilingual speech. *The International Journal of Bilingualism* 3(4): 309–332.

Bentahila, Abdelali & Davies, Eirlys E. 1998. Codeswitching: An unequal partnership? In *Codeswitching Worldwide*, Rodolfo Jacobson (ed.), 25–49. Berlin: Mouton de Gruyter.

Clyne, Michael. 2003. *Dynamics of Language Contact. English and Immigrant Languages* [Cambridge Approaches to Language Contact]. Cambridge: CUP.

Hallap, Valmen. 1960. Mordva-vene segakeelsusest (On Mordvin–Russian bilingualism). *Keel ja kirjandus* 111: 220–223.

Keresztes, László. 2011. *Bevezetés a mordvin nyelvészetbe* (Introduction to Mordvin Linguistics). Debrecen: Debrecen University Press.

Muysken, Pieter. 2000. *Bilingual Speech: A Typology of Code-mixing*. Cambridge: CUP.

Muysken, Pieter. 2007. Mixed codes. In *Handbook of Multilingualism and Multilingual Communication*, Peter Auer & Li Wei (eds.), 315–339. Berlin: Mouton de Gruyter.

Myers-Scotton, Carol. 2001. The matrix language frame model: Development and responses. In *Codeswitching Worldwide II*, Rodolfo Jacobson (ed.), 23–58. Berlin: Mouton de Gruyter.

Myers-Scotton, Carol. 2002. *Contact Linguistics*. Oxford: OUP.

Pineda, David. 2009. Переключение кодов или ‘смешанный лект’ (Code-switching or fused lect]? *Полярный вестник* 12: 26–42.

Pyöli, Raija. 1996. Venäläistyvä aunuksenkarjala: Kielenulkoiset ja -sisäiset indikaattorit kielenvaihtotilanteessa (Olonets Karelian under the Pressure of Russian: Extralinguistic and Linguistic Indicators in a State of Language Shift). Ph.D. dissertation, University of Joensuu.

Raun, Alo. 1988. The Mordvin language. In *The Uralic Languages: Description, History and Foreign Influences* [Handbuch der Orientalistik 8], Denis Sinor (ed.), 96–110. Leiden: Brill.

Russian Census. 2002. ⟨http://www.perepis2002.ru/index.html?id=17⟩ (9 January 2013).

Russian Census. 2010. ⟨http://www.perepis-2010.ru/results_of_the_census/results-inform.php⟩ (9 January 2013).

Sarhimaa, Anneli. 1999. *Syntactic Transfer, Contact-induced Change, and the Evolution of Mixed Codes: Focus on Karelian-Russian Language Alternation*. Helsinki: Finnish Literature Society.

Timberlake, Alan. 1993. Russian. In *The Slavonic Languages*, Bernard Comrie & Greville Corbett (eds), 827–886. London: Routledge.

WALS. 2013. *The World Atlas of Language Structures Online* ⟨http://wals.info/chapter/81⟩ (9 January 2013).

Wertheim, Suzanne. 2003. Linguistic Purism, Language Shift, and Contact-induced Change in Tatar. Ph.D. dissertation, Berkeley CA.
Zabrodskaja, Anastassia. 2009. Russian–Estonian Language Contacts: Grammatical Aspects of Language Use and Change. Ph.D. dissertation, Tallinn University.
Zaicz, Gábor. 1998. Mordva. In *The Uralic languages*, Daniel Abondolo (ed.), 184–218. London: Routledge.

Colonialism, nationalism and language vitality in Azerbaijan

John M. Clifton
SIL International and University of North Dakota

Many of the less-widely-used languages in countries that emerged from the USSR are endangered as a result of Russification. Azerbaijan is home to a dozen less-widely-used languages. Various writers have claimed that most are endangered, although the shift is to Azerbaijani, not Russian. Research conducted from 1998 to 2002, however, found that most of the languages were actively spoken in at least core areas. In this paper I examine factors that led to claims of endangerment. Analyzing the situation in terms of language ecology and its relationship to colonization, I argue that Russian, Azerbaijani, and the less-widely-used languages filled different niches and so did not compete with each other. Since independence, however, shift towards Azerbaijani has accelerated. As a result of nationalism, Azerbaijani and these languages are competing for the same niche. I propose that by expanding the domains of the less-widely-used languages, they can coexist with Azerbaijani.

1. Introduction

Discussions of language endangerment must consider language policies and practices. In the USSR, these were included in the so-called "nationalities question," that is, how a multitude of ethnicities were to be united into a single nation. Grenoble (2003) explicitly examines the role the language policy of the USSR played in the debate regarding the nationalities question. This policy had two goals. The first was to protect the rights of non-Russian-speaking communities to use their traditional languages in education and other domains, while the second was to encourage every citizen to learn Russian. Bilingual schools, where students began in the local language and then moved to Russian, represented a way to meet both goals. Over time, however, the goal of protecting the use of traditional languages became secondary under a stronger push for Russification. Grenoble notes that as the use of Russian spread, many of the less-widely-spoken languages became highly endangered. Neroznak (1994) documents 63 endangered languages just in Russia; as Grenoble (2003: 197) notes, this list includes all the languages spoken in the

Far East. And language endangerment associated with Russification did not only affect languages in Russia. The list of threatened "national cultures" presented by Vahtre and Viikberg (1991) includes endangered languages from other post-Soviet countries including Azerbaijan, Georgia, Tajikistan and Ukraine.

Of course, language endangerment is not limited to the USSR. It is generally accepted that multilingualism and language shift have led to language endangerment and death in many areas of the world. These effects have not, however, been universal. Over 15 years ago, Crowley (1995) claimed that traditional languages in Vanuatu had greater vitality than would be expected on the basis of claims by Mühlhäusler (1987) and Dixon (1991). My own research indicated that the Kaki Ae language of Papua New Guinea existed in a state of stable multilingualism in spite of the fact that it was unrelated to any surrounding languages and was spoken by only about 300 people (Clifton 1997). These instances of unexpected language vitality are significant since they can shed light on the causes of language endangerment.

In this paper I examine the relationship between multilingualism, proficiency in the language of wider communication, diglossia, and language shift in Azerbaijan, one of the countries that emerged from the break-up of the Soviet Union. I show that while most speakers of less-widely-used languages are multilingual, and proficiency in Azerbaijani was high even in Soviet times, large-scale shift to Azerbaijani is a fairly recent phenomenon. Instead of shift, widespread diglossia has been the rule. Finally, I explain the spread of diglossia during Soviet times and language shift more recently in terms of the model of language ecology developed by Mufwene (2008). The application of Mufwene's framework to the linguistic situation in Azerbaijan is especially important because it provides reasons for the stability of less-widely-spoken languages that have been missing in past studies. At the same time, the situation in Azerbaijan forces us to reevaluate the types of colonization that form the basis of Mufwene's analyses elsewhere.

With approximately 8.8 million people, Azerbaijan is located in the southern Caucasus as shown in Figure 1.[1]

The linguistic situation in Azerbaijan is complex. The state language is Azerbaijani, a Turkic language. In addition to Azerbaijani, Azerbaijan is traditionally home to six Indo-European languages (Armenian[2] [hye],[3] Khalaj [kjf],

1. Map is from *The world factbook* (2011).

2. While Armenian is a traditional language of Azerbaijan, most speakers of Armenian have either left Azerbaijan as a result of the Nagorno-Karabakh conflict, or live in the contested region of Nagorno-Karabakh.

3. The three-letter codes are the ISO-639-3 unique language identifiers.

Figure 1. Azerbaijan in context

Kurdish [kmr], Mountain Jewish or Judeo-Tat [jdt], Talysh [tly], Tat or Muslim Tat [ttt]), eight North Caucasian languages (Avar [ava], Budukh [bdk], Khinalug [kjj], Kryts [kry], Lezgi [lez], Rutul [rut], Tsakhur [tkr], Udi [udi]), and one Kartvelian language (Inghiloi or Georgian [kat]) as shown in Figure 2.[4]

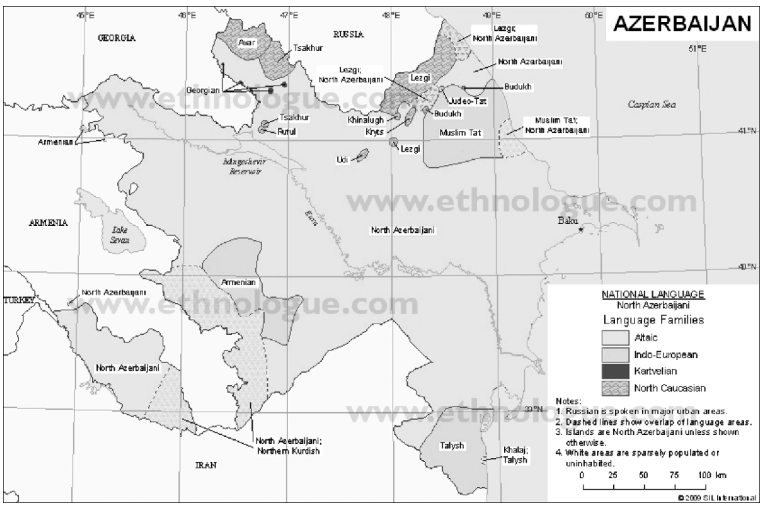

Figure 2. Languages of Azerbaijan

In addition to these indigenous languages, Russian has played and continues to play a role in society and government. This mosaic of languages has resulted

4. Map is from Lewis (2009).

in widespread multilingualism. As a result, Azerbaijan provides a good testing ground for theories regarding the relationship between multilingualism and language shift.

According to the 2009 Azerbaijani census, over 91% of the population is ethnic Azerbaijani, while the 1999 census indicates that over 90% speak Azerbaijani.[5] The official population of most of the other indigenous ethnic communities ranges from as few as 1000 to as many as 180,000. Although the rights of speakers of less-widely-used languages are enshrined in the 1992 presidential decree "On the Protection of the Rights and Liberties, and Development of Languages and Cultures of Ethnic Minorities Living in the Territory of Azerbaijan Republic," it has been claimed that most are endangered. Vahtre and Viikberg (1991) include ten languages from Azerbaijan in *The Red Book of the Peoples of the Russian Empire*. UNESCO's *Atlas of the World's Languages in Danger* categorizes three of the languages (Avar, Lezgi, Talysh) as vulnerable, three (Mountain Jewish, Rutul, Tsakhur) as definitely endangered, and five (Budukh, Khinalug, Kryz, Tat, Udi) as severely endangered (Moseley 2010).

Against the backdrop of these claims, field research conducted in Azerbaijan from 1998 to 2002 indicated that most heritage languages were still vital in at least a significant core region. My purpose in this paper is to investigate why it has been assumed that most languages in Azerbaijan are endangered, possible reasons for their unexpected vitality, and the prognosis for their long-term survival. I begin in Section 2 by focusing on aspects of Soviet language policies that would support the conclusions of Vahtre and Viikberg (1991) and Moseley (2010) that most of the languages of Azerbaijan are endangered. Next, in Section 3 I summarize our findings for a number of languages spoken in Azerbaijan, justifying the claim that they show unexpected vitality. In Section 4 I examine the implications of Mufwene's claim that many cases of language shift and death can be explained by patterns of colonization (Mufwene 2008). I show in that section that although Russian colonization does not fit neatly into either of the two major patterns outlined by Mufwene, more general principles of language ecology underlying his analysis do in fact predict the vitality seen in our research. In Section 5 I examine the current effects of nationalization, showing it is triggering effects comparable to those Mufwene describes as resulting from settlement colonization. I conclude by outlining possible responses in Section 6.

---

5. The 2009 census includes information on ethnicity but not language. Data for the 2009 census is from the State Statistical Committee of the Republic of Azerbaijan (2012); data from the 1999 census is from *The world factbook* (2011).

2. Soviet language policies in Azerbaijan

In 1920, Azerbaijan became an autonomous republic in the Soviet Union. While the official policy of the USSR was that all languages and cultures should have equal opportunities for development and support, Azerbaijani (with Armenian and Georgian) were recognized as official languages in their respective republics by the Soviet state. This elevated status of Azerbaijani was reflected in the educational policies. Children in Azerbaijani-speaking communities did not need to go to schools where Russian was the language of instruction: Communities could choose whether the local schools would use Russian or Azerbaijani as the language of instruction. In mixed communities, local schools had two separate 'sectors,' a Russian-language sector and an Azerbaijani-language sector. Furthermore, while there were Russian-only schools for Russian communities in the Russian SSR, Azerbaijani was a required subject in all Russian-language schools in Azerbaijan.[6] At the same time, Russian was taught as a subject in schools where Azerbaijani was the language of instruction (Garibova & Asgarova 2009: 195).

In spite of its official status, however, Azerbaijani was subordinate to Russian. While Russian had no official status, it was the de facto language of government. Furthermore, Russian-language classes were better supplied than Azerbaijani classes even in schools where Azerbaijani was the language of instruction. In urban areas, children from middle class families tended to attend Russian language schools (Arutiunov 1998: 105). The relative status of Russian and Azerbaijani is reflected in the fact that in 1980 seventy percent of students at the Azerbaijan State University studied in Azerbaijani, while all classes at the more prestigious Oil and Gas Institute were in Russian (Altstadt 1992). The result was that advancement in government, education, research and technical professions was dependent on fluency in Russian (Garibova 2009: 13).

Although Azerbaijani was subordinate to Russian, less-widely-used languages were even further marginalized. As Grenoble (2003: 111) notes, while Azerbaijani competed only with Russian, these languages had to compete with both Russian and Azerbaijani. In Azerbaijan, the guarantee that all children should have access to education in their mother tongue was ignored in the case of these languages (Grenoble 2003: 111–112). Since there was room in the curriculum for only two languages, and given that both Azerbaijani and Russian had to be included in the curriculum, the heritage language was left out. While it might be possible to claim that this was due to a lack of materials, the situation in languages like Lezgi and

6. This was not true in those few schools where Georgian was the medium of instruction. Russian was the second language at these schools; Azerbaijani was not taught as a subject.

Avar in Dagestan, a neighboring republic that was a part of Russia, shows this is not the case. Since none of the indigenous Dagestani languages had any official status, not one of them had to be used in all schools. Because of this, different indigenous languages could serve as either the language of instruction (in lower grades) or as the second language in schools in Dagestan. Materials were created in the Lezgi and Avar languages for use in their respective communities in Dagestan, and Lezgi/Russian schools and Avar/Russian schools were common in these communities. But even though Lezgi pedagogical materials existed, Lezgi children living in Azerbaijan had to study in Russian/Azerbaijani or Azerbaijani/Russian schools (Kosvena 1960).

While languages like Lezgi, Talysh, Udi and Khinalug had no official status in Azerbaijan during the Soviet era, the fact that significant numbers of people in Azerbaijan did not speak Azerbaijani as their first language was used by Soviet officials in Moscow to promote Russian as a general lingua franca. Changing Latin-based orthographies to Cyrillic-based orthographies in the 1930s was another way of promoting Russian norms (Grenoble 2003: 115).

Given the complex relationships between Russian, Azerbaijani, and the less-widely-spoken languages, and the concerted effort of the central government to promote Russification, it would not have been surprising to find widespread shift from all the indigenous languages of Azerbaijan, including Azerbaijani, to Russian. This is exactly what happened in Central Asian countries like Kazakhstan, where people in the cities gradually became Russian monolinguals (Fierman 2005). In Azerbaijan, however, this did not happen. Russian never became established in the rural areas. Azerbaijani was the medium of instruction in most schools in these areas, and the inhabitants of these communities rarely spoke Russian. Garibova and Asgarova (2009: 195) argue that constant movement between the urban and rural regions acted to maintain the primary role of Azerbaijani in the republic. As a result, even Russian-speaking Azerbaijanis holding positions in government or education were bilingual in Azerbaijani (although they were generally more fluent in Russian).

While there was no widespread shift from Azerbaijani or the less-widely-spoken languages to Russian, and the role of Russian was limited, many researchers reported that members of communities speaking less-widely-used languages exhibited high levels of proficiency in Azerbaijani. This phenomenon led to the claim that these communities were undergoing language shift from the heritage languages to Azerbaijani. The claim was not based on actual research in the communities to determine what the actual levels of Azerbaijani proficiency were, however, or whether the expected shift was actually occurring. It was simply assumed that shift must be occurring given the extensive bilingualism. In the next section I present data indicating that while Azerbaijani proficiency was indeed high, widespread shift to Azerbaijani was not occurring.

3. Language vitality assessed

From 1998 to 2002 I led a research team investigating the vitality of eleven indigenous languages in Azerbaijan. For each language, we visited a representative sample of villages, interviewing government officials, educators, religious leaders, medical personnel, and groups of village people. In some locations we also used Sentence Repetition Tests in which respondents repeat a series of graded sentences in the test language (Radloff 1991),[7] and Perceived Benefits questionnaires where respondents indicate how important various languages are perceived to be for various purposes (Karan 2011). Our goals were to determine levels of proficiency[8] in the traditional languages, Azerbaijani and Russian; to determine patterns of language use; and to determine attitudes towards various languages. In Section 3.1 I present our findings regarding two Iranian languages, Talysh and Tat. Next I present the findings regarding Budukh, Kryz and Khinalug, three North Caucasian languages located in the Shahdagh Mountains. Finally, in Section 3.3 I discuss the situation in Lezgi and Udi, two communities where the Russian language figures more prominently.

3.1 Talysh and Tat

Talysh and Tat are two of the four Iranian languages indigenous to Azerbaijan. According to the 2009 census there are 120,000 Talysh and 25,200 Tat in Azerbaijan. The two languages are physically separated from each other: Talysh, spoken in the south of the country, is indigenous to Azerbaijan and Iran, while the Tat language is spoken in northeastern Azerbaijan. The categorization of Talysh as "vulnerable" and Tat as "severely endangered" (Moseley 2010) reflects the differences in population size: There are considerably more Talysh than Tat. The claimed

7. A Sentence Repetition Test is calibrated by having non-native speakers of independently-determined proficiency repeat each potential test sentence. On this basis, sentences are chosen for the final test by their ability to differentiate consistently between different proficiency levels. The overall calibration of the test involves correlating the performance of the original pool of non-native speakers with their independently-determined levels of proficiency. Radloff (1991) demonstrates that the resulting SRT is valid when used with a sufficiently large group of test subjects. Furthermore, the results of our use of the SRT correlated with the subjective evaluations given by native speakers of Azerbaijani and test subjects.

8. Most determinations of proficiency were made through self-reporting. However, these reports were corroborated by a number of other methods. As indicated, a Sentence Repetition Test was used in some communities to determine proficiency more objectively. Secondly, teachers and medical personnel, many of whom were native speakers of Azerbaijani, were asked to comment on levels of proficiency. Thirdly, the researchers interacted with participants in Russian and/or Azerbaijani, and were able to gauge proficiency during these interactions.

difference in endangerment reflects previous reports of proficiency in Azerbaijani: While Rastorgueva (1979) claims that the Talysh had high levels of proficiency in Azerbaijani, Grjunberg and Davidova (1982) claim that the Tat considered Tat and Azerbaijani as equal first languages.

Clifton, Deckinga, et al. (2003) and Clifton, Tiessen, et al. (2002) verify claims that both groups exhibited high levels of proficiency in Azerbaijani as measured by an Azerbaijani Sentence Proficiency Test (Radloff 1991) in which people are asked to repeat a graded set of sentences in Azerbaijani. In the most isolated Talysh village visited, 60% of the people tested exhibited near-native proficiency. The level of proficiency was higher in other communities: 75% of the people tested in other Talysh villages and 84% of the people tested in Tat villages exhibited near-native proficiency.

While proficiency in Azerbaijani was uniformly high, levels of proficiency in Talysh and Tat were more varied. In ethnically-mixed villages, fluency was reported to be low among individuals under the age of 45. In homogenous communities, however, self-reported data indicated that people of all ages and both genders had high levels of oral proficiency in the traditional languages. Isolation also played a role in fluency, as children in more isolated homogenous villages were reported to acquire the traditional language by the age of six, while children in less isolated homogenous villages reportedly took until 15 to reach fluency. At least in the Tat community, a final factor in fluency was economic opportunity. Children in villages with little economic opportunity were encouraged to become more fluent in Azerbaijan, and showed less fluency in the traditional language. One explanation for this is that parents in these villages were promoting the use of Azerbijani as a way to improve their lot in life even if this negatively affected fluency in the heritage language. This explanation is supported by reports from officials that people are moving out of the village to find jobs in Azerbaijani-speaking areas, and that they move back if they cannot find a job outside the village.

In general, Azerbaijani existed in a diglossic relationship with both Tat and Talysh: Azerbaijani was used in formal situations including education, government and the media, and the traditional languages were used in informal situations including the home and local community. Our research indicates, however, that the factors that determine fluency were also relevant to analysis of language use. While diglossic relationships existed in most communities, the use of Tat and Talysh was greater in homogenous communities. In ethnically-mixed locations, Azerbaijani was used more widely with Tat and Talysh playing a secondary role. Isolation was also a factor, as people in more isolated villages used Tat and Talysh more commonly in informal situations than did people in less isolated situations.

At the time of this research, then, high levels of proficiency in Azerbaijani had resulted in diglossia rather than in an overall shift from Tat and Talysh to Azerbaijani. While individuals under 45 in ethnically-mixed villages had little to no proficiency in the traditional language, there was a good-sized core in each language community where proficiency in the heritage language remained high.

3.2 The Shahdagh languages

The three Shahdagh languages, Budukh, Kryz and Khinalug, are North Caucasian languages traditionally spoken near Mount Shahdagh in northeastern Azerbaijan. According to the 2009 census, Khinalug is the largest of the three communities with a population of 2,200, less than ten percent of the population of Tat. Budukh and Kryz have even smaller populations and are not listed separately in the census data. Instead, they are included with "other nationalities." There has been considerable interaction between the three communities for generations. The primary occupation in all three communities is sheep herding, and during the Soviet era they made up a single collective farm, pasturing their sheep together. Qübatov (1991) claims that members of all three communities were so fluent in Azerbaijani that they could not determine whether they were more fluent in their heritage language or Azerbaijani.

Our research, summarized by Clifton (2009), confirms that Azerbaijani proficiency was generally high in all three communities, especially among men under the age of 55. The use of common pasture lands even after the dissolution of the collective has reinforced the use of Azerbaijani as a lingua franca. We also confirm claims that large numbers of residents of all the mountain villages have migrated to the plains where they live in proximity with native speakers of Azerbaijani. Interviewees indicated this movement had its roots generations ago in the use of summer pastures in the mountains and winter pastures in the lowlands.

In light of these factors, it is not surprising that Moseley (2010) classifies all three languages as "severely endangered." But our research indicates that the situation is much more complex in the areas of proficiency in the traditional languages and language use, with major differences between the three groups. These differences resulted in different levels of endangerment.

The Budukh language was gravely endangered. Migration to ethnically-mixed villages in the plains was higher in Budukh than in either Khinalug or Kryts. There were only 43 households in Buduq,[9] the main mountain village, in 2001, down

9. I use standard Azerbaijani spellings for village names.

from 98 households ten years before. The school only went through grade 9, and there were no medical facilities in the village. The plains villages, on the other hand, were thriving.

Proficiency in Budukh was especially low in the plains villages. It was reported that young people and children in the plains villages did not know Budukh. Instead, they interacted in Azerbaijani. Even among adults who knew Budukh, Azerbaijani was the usual language of communication in the village. The vitality of Budukh was somewhat better in the core village of Buduq. But even there, both Azerbaijani and Budukh were used in the home, and it was estimated that 60% of children knew Azerbaijani before entering school.

While the Kryz language is very closely related to Budukh linguistically, its sociolinguistic situation was quite different. The migration from mountain to plains villages was not as widespread for Kryz as for Budukh. Furthermore, one cluster of plains villages was homogenously Kryz. So while there had been movement to ethnically-mixed plains villages, this represented a minority of the overall community and the migration had less impact on the vitality of Kryz than on the vitality of Budukh.

While men under 55 living in Kryz mountain villages were reported to speak Azerbaijani well, the same was not true for women and men over 55. Since residents of mountain villages have little contact with first-language Azerbaijani speakers, fluency in Azerbaijani is tied to schooling. But educational levels were reported to be low. Only one of the five mountain villages has a school through grade 9, and less then half of the few students who do finish grade 9 go for further education. Residents said that interest in education is low since the economy was based on sheep husbandry. While proficiency in Azerbaijani is relatively low, proficiency in Kryz was reported to be high, with the language in daily use in the home and community (Clifton 2009: 39).

Proficiency in and use of Kryz was reported to be lower in plains villages than in mountain villages. While most people over 30 speak Kryz as their first language, a significant number are married to non-Kryz and so use Azerbaijani in the home. Reports were that in the plains villages no more than thirty percent of Kryz between the ages of 18 and 30 could speak Kryz well. As a result, respondents estimated that only thirty to forty percent of children are growing up in homes where Kryz was still being used even occasionally. The children in most plains villages were reported to be more fluent in Azerbaijani than in Kryz (Clifton 2009: 39–40). Overall, then, Kryz was being actively used in all of the mountain villages, and in some of the lowlands villages.

Finally, the Khinalug community exhibited diglossic patterns similar to those shown by the Tat and Talysh communities. Up to eighty percent of ethnic Khi-

nalug lived in one of two mountain villages. There was a school with grades 1–11 and a clinic in Xınalıq, the main mountain village. Even though travel to and from Xınalıq was generally difficult, and the village was completely cut off for much of the winter, migration from the mountains to the villages was considerably less than it was in Budukh or Kryz.

Our research indicated that there was no major shift by ethnic Khinalug from the traditional language to Azerbaijani. While Azerbaijani was being used with outsiders and in education, Khinalug continued to be used in the home and village. The Khinalug highly valued their language, and wanted their children to learn both Khinalug and Azerbaijani well. In interviews, individuals commonly indicated that mothers should teach their children Khinalug before Azerbaijani. Khinalug seemed to be undergoing less shift than either Budukh or Kryz.

The situation in the Shahdagh languages can be summarized as follows.

- Budukh was gravely endangered with speakers moving to the plains and shifting to Azerbaijani even in the mountain villages.
- Kryz was being actively used in mountain villages even as shift to Azerbaijani was occurring in many plains villages.
- Khinalug exhibited a diglossic situation where Azerbaijani was used with outsiders and in education, and Khinalug was used in the home and community.

3.3 Udi and Lezgi

This quick overview of the sociolinguistic situation of a few of the languages in Azerbaijan illustrates the point made in Section 2 regarding the relative prominence of Azerbaijani and Russian. In spite of attempts at Russification, Azerbaijani was generally more important than Russian in the life of communities speaking less-widely-used languages. It was more common as a second language, and when language shift occurred it was generally towards Azerbaijani, not Russian. There were, however, communities in which Russian could be expected to play a more prominent role. This is especially clear when we look at data from the 1989 Soviet census on second language use. While this census does not include many of the languages of Azerbaijan, it does include information for Talysh, Tsakhur, Udi and Lezgi. Figure 3 shows what language people reported using as a second language in this census as reported in Trosterud (1998). The three options were Russian, Other, and Neither. Monolinguals reported 'Neither;' given the fact that Azerbaijani is the most commonly used language of wider communication, we can assume that most of those responding with 'Other' in Azerbaijan spoke Azerbaijani as a second language.

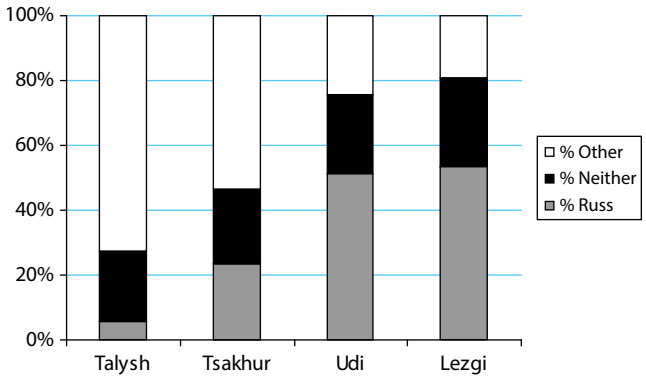

Figure 3. Second language use for Talysh, Tsakhur, Udi and Lezgi

Over 70% of ethnic Talysh reportedly spoke Azerbaijani as their second language, as opposed to less than 6% who spoke Russian. A larger percentage of Tsakhur spoke Russian as a second language, but this must be tempered by the fact that almost 25% of Tsakhur lived in Russia, not Azerbaijan. The situation among the Lezgi and Udi is interesting. The high level of Russian as a second language among Lezgi is undoubtedly due to the fact that the majority of Lezgi were reported to live in Russia. The high level of Russian as a second language among the Udi, however, cannot be due to place of residence, since most Udi lived in Azerbaijan. A more reasonable explanation is that they are the only ethnic group in Azerbaijan who are Christian, and as such have traditionally viewed themselves as aligned with the Russians. This is corroborated by responses of Udi speakers to questions regarding language attitudes during sociolinguistic research. When asked if it was all right for an Udi to marry an Azerbaijani or a Russian speaker, respondents who indicated that it would be acceptable to marry an Azerbaijani speaker stipulated that the Azerbaijani speaker should be a Christian. There was no such stipulation regarding Russian speakers (Clifton, Clifton, et al. 2002: 115).

Since, according to the census, more than half of the Lezgi lived in Russia, it is significant that the level of Russian as a second language among the Lezgi is only slightly more than 50% and the level of Azerbaijani as a second language is almost 25%. Given that there was interaction between the Lezgi communities across the Azerbaijan/Dagestan border and the fact that a majority of Lezgi children attended schools in which Russian is the language of instruction, it would not have been surprising if a larger number of Lezgi claimed Russian as a second language. But this census data mirrors our own findings reported in Clifton, Lucht et al. (2002). Residents of Lezgi villages in Azerbaijan we interviewed consistently indicated that the majority of Lezgi in Azerbaijan had higher or equivalent proficiency in

Azerbaijani as compared to Russian. This was true for both men and women of all age groups.

The situation was even more interesting among the Udi as detailed in Clifton, Clifton, et al. (2002). As expected, a majority of respondents indicated that they spoke Udi more fluently than either Russian or Azerbaijani. A separate group of interviewees was given a list of ten situations and asked whether they would be able to handle each one in Udi, Azerbaijani and Russian. While they all indicated they would be able to deal with the first six situations in all three languages, a majority said they could only handle the final four situations in Russian. These four situations included describing how they spent their free time, describing their employment or studies, describing their hopes for the future, and giving an opinion. Assuming the respondents were fluent in Udi, these responses imply they are also fluent in Russian. These results were not confirmed by a Russian Sentence Repetition Test (SRT) where Udi speakers were asked to repeat a series of fifteen graded sentences in Russian (Radloff 1991). Only two of 46 Udi speakers who took the Russian SRT scored at native speaker level. Over half (26) scored at a level of "good, general proficiency" or lower, hardly native speaker level.

While we found in most groups that self-assessment responses were close to the objective measures of proficiency, this was obviously not the case for the Udi self-assessment of proficiency in Russian. An explanation of this situation draws on a number of factors. Although most Udi have attended schools in which Russian is the medium of instruction since 1953, responses to questions regarding language use indicated that most do not have to use it on a regular basis. Furthermore, most have little if any actual interaction with native speakers of Russian. Instead of comparing their command of Russian with that of native speakers, they are comparing it with that of their Azerbaijani neighbors. Simultaneously, they are comparing their command of Azerbaijani with that of their Azerbaijani neighbors, and that makes it appear that their proficiency in Azerbaijani is lower than it actually is. Another factor is that the Russian spoken in Azerbaijan is different from that spoken in Moscow;[10] the Russian SRT tests proficiency in the Russian dialect. Low scores on the Russian SRT indicate that while the speakers may well be able to use the local dialect of Russian, they will have problems when confronted with standard Russian. Anecdotal evidence from Russian speakers who have interacted with Udi speakers corroborate this expectation.

10. For example, Russian has masculine and feminine pronouns, and a basic verbal contrast between perfective and imperfective. Neither Azerbaijani nor Udi have these distinctions, and these distinctions are also lost in Azerbaijani Russian.

One final observation concerning Udi is that it was still being used extensively even though a third of the residents of Nic, the primary Udi village, are ethnic Azerbaijani. Over two-thirds of respondents indicated they use Udi in the workplace. Udi is the language of the home and the village. Respondents unanimously said that mothers should speak Udi to their children and that children should learn Udi before Russian or Azerbaijani. The commitment of the Udi community to Russian, then, did not in any way lessen their commitment to the Udi language.

4. Colonization patterns and language vitality

Two important points follow from the case studies presented in Section 3. First, the less-widely-used languages of Azerbaijan showed more vitality than was expected on the basis of previous studies. While at least one (Budukh) was gravely endangered, several seemed to have developed a stable diglossia with Azerbaijani, at least in core locations. Second, the situations found in Lezgi and Udi indicate that attempts by Soviet authorities to Russify communities speaking less-widely-used languages do not seem to have worked even when there were factors that favorably predisposed the communities to using Russian. While Russification has led to serious endangerment of less-widely-used languages in the Far East of Russia, it did not have a similar effect in Azerbaijan.

A possible explanation for this situation is provided by correlating language shift with patterns of colonization. Mufwene (2008) highlights how different language shift patterns are in Africa as compared with those in the Americas and Australia. Overall, language endangerment has not been as extensive in Africa as it has been in the Americas and Australia. In the Americas and Australia, shift to the colonial languages led to endangerment and even death of many of the indigenous languages. In Africa, on the other hand, language shift (when it occurred) has most commonly involved shift from a traditional language to a local lingua franca, not to the colonial language.

Mufwene accounts for these differences in terms of language ecology, that is, the different functions played by competing languages within the linguistic environment, arguing that different patterns of colonization result in different relationships between the colonial and indigenous languages within the ecological system. Mufwene claims that colonization usually began as trade colonization, in which the colonizer and colonized simply traded with each other on a roughly equal footing. Contact was sporadic, and communication was accomplished through a few local interpreters. While trade colonization sometimes resulted in pidgins, it did not result in language shift since the languages of the colonizer and the colonized did not compete.

In most places, trade colonization quickly developed into either settlement colonization or exploitation colonization. In settlement colonization as found in the Americas and Australia, large numbers of settlers came from the colonizer to the colony, intending for their descendents to remain in the colony. Ultimately, they overwhelmed the indigenous inhabitants. The colonial language filled the same socio-economic niche as had been filled by traditional languages. In pre-colonial days, the indigenous languages had been used for all economic activity. In settlement colonies, the colonists both maintained their own language and controlled economic development. Any of the indigenous people who wanted to participate in this new economy had to use the language of the colonizer. This conflict between the languages of the colonizer and colonized has led to profound language endangerment in the Americas and Australia.

In exploitation colonization as was practiced in Africa, on the other hand, immigrants from the colonizer did not intend to remain in the colony permanently. Instead, the colonizer sent a relatively small number of administrators who planned to remain in the colony for a finite period of time before being replaced by new administrators. The administrators, in turn, worked through bilingual indigenous people to mobilize the local communities on behalf of the colonizer. The point of exploitation colonization was to extract resources from the colony, not to use the colony as a new home. In exploitation colonization, the languages of the colonizer and colonized filled separate socio-economic niches. Representatives from the colonizer did not need to learn the indigenous languages, and outside of the translators, the indigenous people did not need to learn the colonizer's language. The end result of exploitation colonization is a core of bilingual indigenous people.

Russian colonization cannot be neatly classified as either settlement colonization or exploitation colonization. After centuries of conflict between the Iranian Safavids, Ottomans and Russians, Catherine established a permanent Russian presence in Azerbaijan in the 1790s. Conflicts between Iran and Russia continued for some time, but Iran officially recognized Russian sovereignty over the region in treaties of 1813 and 1828.

While Azerbaijan was already producing oil, and a main objective for Soviet control of Azerbaijan was to gain access to Azerbaijani oil (Landau & Kellner-Heinkele 2001: 15), this was not a factor for Catherine since commercial development of oil in the region did not begin until the 1870s. For Catherine, the principle purpose of colonization was expansion and prestige (Altstadt 1992); it was a staging point for expansion (Swietochowski 1995: 12–13).

As is common in settlement colonization, long-term Russian settlers migrated to Azerbaijan. This began in tsarist times when Russians in the military who served in the Caucasus settled there (de Waal 2010). Immigration from Russia

continued throughout both the tsarist and Soviet periods. After 1945, most of this immigration was voluntary, with the new settlers considering Azerbaijan part of Russia (Landau & Kellner-Heinkele 2001: 35–36). As in the Americas and Australia, these settlers did not learn the local language or culture. Only 14.4% of Russians in Azerbaijan knew Azerbaijani at the time of the break-up of the Soviet Union, even though many were from families who had lived in Azerbaijan for generations (Landau & Kellner-Heinkele 2001: 17).[11]

Other aspects of Russian colonial rule were characteristic of exploitation colonization. Russians constituted a minority of settlers in Azerbaijan during the Tsarist period; the majority of immigrants were Armenians (Swietochowski 1995: 10–11). While immigrants to the Americas and Australia settled in urban and rural areas, most Russians in Azerbaijan lived in towns and cities. Over 75% of all Russians lived in Baku, the capital (Landau & Kellner-Heinkele 2001: 44), while 95% lived in towns or cities (Landau & Kellner-Heinkele 2001: 36). Even those Russians in rural areas lived in homogenous Russian villages that were separated from Azerbaijani villages. Most Russians in Azerbaijan were either bureaucrats or held specialized, technical jobs (Landau & Kellner-Heinkele 2001: 37). While Russian currency was used, Russia made little economic investment into Azerbaijan through the mid 19th century, seeing it primarily as a source of temperate-zone crops (Swietochowski 1995: 17). Finally, the colonial administrations were small as is common in exploitation colonization. In fact, de Waal (2010) claims that Tsarist Russia lost control over the Caucasus in the early 20th century because the administration was too small to intervene in local conflicts.

In terms of language ecology, Russian and the indigenous languages filled different niches. The Russian colonial administration controlled little of the socio-economic life of ordinary Azerbaijanis; Azerbaijanis who were not involved in government or education did not need to learn Russian to fully participate in their traditional economic activities. Furthermore, there were a limited number of jobs available for the Russian-speaking Azerbaijani. Therefore, there was little incentive for the typical Azerbaijani to learn Russian. My conclusion is that the key distinction between settlement and exploitation colonization is whether the colonizer and indigenous languages fill the same niche, not whether there are substantial numbers of long-term settlers. In Azerbaijan, Russian and the indigenous languages (including Azerbaijani) did not fill the same niche, and so there was little shift to Russian.

---

11. This figure was, however, significantly higher than comparable figures in Central Asia where one to five percent of Russian immigrants learned the titular languages (Landau & Kellner-Heinkele 2001: 37).

5. Nationalism and language vitality

The ecological paradigm that explains the lack of shift to Russian can also be used to explain the general lack of shift to Azerbaijani by speakers of less-widely-used languages. Azerbaijani and these languages did not fill the same socio-economic niche. Azerbaijani was needed for education, for local government, and for communication with people from neighboring language groups. But it was not needed for ordinary life in the home and community. Just as Russian was not needed for the traditional economic activities, Azerbaijani was not needed either. This resulted in wide-spread diglossia, but little shift to Azerbaijani.

The shift to Azerbaijani by speakers in the core mountain villages of Budukh (see Section 3.2), can also be explained in terms of language ecology. The Budukh community is located in the midst of one of the most linguistically diverse regions in Azerbaijan. According to Rafik Abdullayev, president of the Buduq Society (personal communication), Buduq was one of the largest villages in the area with over 500 houses in the 19th century. For this reason, the Soviets made Buduq village the regional center. A full school was established there which provided education for children in Buduq and the surrounding villages, while teachers from the Budukh community opened schools in other language areas. Medical facilities in Buduq village served the entire region. These changes would have resulted in a situation in which members of the Budukh community were no longer limited to the traditional economic activities. To become part of this new economy, however, it would have been necessary to know Azerbaijani. At that point Azerbaijani and Budukh would have been filling the same socio-economic niche. Ultimately this would have led to precisely the situation we found where the majority of speakers had shifted to Azerbaijani. Even though Buduq village was no longer the regional center, and the schools and medical center had been downgraded or closed, the shift to Azerbaijani could not be reversed.

While our research found that most of the less-widely-spoken languages were not being abandoned, there was evidence that the situation was changing. As noted in Section 3.1. proficiency in the heritage language was lower in Tat communities with little economic development. In many communities, the desire for children to do well in school has resulted in parents speaking only Azerbaijani with school-aged children; in some communities this has extended to pre-school children (Clifton, Tiessen, et al. 2003:87). While we were told that this will not have long-term effects since children will learn Talysh from playmates and grandparents, this definitely represents a recent change.

More recently, I have received informal reports that the shift towards Azerbaijani has accelerated, both by members of communities that traditionally spoke less-widely-used languages and by ethnic Azerbaijanis. This shift seems

to be tied directly to the gaining of independence tied to the break-up of the USSR. Landau and Kellner-Heinkele (2001:44) note that Russians remaining in Azerbaijan after independence complained about the rate of transition to Azerbaijani in the urban areas. We noticed the same phenomenon: Acquaintances who spoke only Russian when we first visited in 1995 were speaking nothing but Azerbaijani by 2001.

Post-independence shifts like this could be the results of overt attempts by the government to mandate language use. Arutiunov (1998: 106) claims this is what has happened in Azerbaijan; that governmental policies are designed to displace both Russian and less-widely-spoken languages by Azerbaijani, and to force the minority groups to assimilate. This was true to at least some extent of the government led by the Azerbaijan Popular Front (APF) in 1992–93 (Marquardt 2010, 2011). This government, led by Abulfez Elchibey, was outspokenly pan-Turkic and anti-Russian. There was a strong attempt during this period to 'Turkeyize' Azerbaijan both culturally and linguistically.

Marquardt (2010, 2011), however, argues that under the subsequent administrations of Heydar and Ilham Aliev, government policies towards language have been mostly symbolic. On the one hand, Russian-language schools have been allowed to continue, while on the other, materials have been produced in less-widely-spoken languages and the linguistic rights of these communities has been affirmed. While the Aliev governments have emphasized their links to Turkey, they have simultaneously emphasized their differences. This is evident even in the implementation of the shift from Cyrillic to Latin orthographies: While the Azerbaijani Latin orthography is almost identical to the Turkish orthography, the presence of the unique letter ⟨ə⟩ in Azerbaijani has been emphasized.

My observations agree with Marquardt's. The recent acceleration towards Azerbaijani is not due to governmental policies, but to changes in the ecology of language. Since independence, Azerbaijani has not had to share official functions with Russian. The role of Azerbaijani in government, education and the economy has become increasingly important. Separatist movements in the Lezgi and Talysh communities in the early 1990s also seem to play a role, as members of those communities are eager to display their commitment to the new nation, and use of the state language is one way to demonstrate that. As Azerbaijani has taken on an increased range of functions, especially in the socio-economic realm, it has entered niches traditionally belonging to less-widely-used languages. Mufwene (2008: 229) notes that in cases of settlement colonization, one of the colonial languages gradually extends its domination until it is the exclusive language of the nation. At that point, indigenous people as well as new immigrants shift to this language. In Azerbaijan, nationalism is playing a role similar to that traditionally played by settlement colonization. Once Azerbaijani fills the niches that have

traditionally been filled by less-widely-spoken languages, the members of these communities find themselves under great pressure to shift to Azerbaijani.

6. Responses

When fashioning responses to language shift, it is first important to determine the attitudes of those affected. The overwhelming sentiment among the speakers of less-widely-spoken languages we interviewed was that they wanted their children to be fully bilingual in the heritage language and Azerbaijani. They were proud of their language, and saw it as a marker of their group identity. The shift we observed, then, was not a shift away from the traditional language, but a shift to Azerbaijani as the two languages increasingly filled the same socio-economic niche. Responses to this situation, then, need to find ways to expand the domains filled by the less-widely-spoken languages.

Auger (2011a, 2011b) presents an instructive example in the French regional language Picard. The factory workers and peasants who traditionally spoke Picard are increasingly shifting to French. Recent surveys, however, indicate that Picard has not died. Instead, it is being spoken and written by professionals and intellectuals. The niche for Picard has shifted, keeping the language alive.

In the case of Azerbaijan, the rights of speakers of less-widely-spoken languages enjoy official government protection. Furthermore, there are members of many of these communities who are professionals in education, government or industry who are interested in promoting their languages. For example, a professional from the petroleum industry has compiled a Tat dictionary (Ağacamal & Məmmədxan 2007), while a researcher at the Azerbaijani Academy of Sciences compiled a dictionary of Talysh (Mammedov 2006). A teacher in the Khinalug school has begun work on a Khinalug dictionary.

In the mid-1990s the government funded publication of primers in a number of languages including Tat, Talysh, and Udi. Unfortunately, we discovered during our research that distribution of the primers has been problematic, and many teachers have not been trained in how to use them. Some of them are also in need of revision. There is interest in establishing classes in vernacular literacy in the community schools, but no one to provide the necessary support.

The question arises, then, as how to assist these individuals and communities in their efforts to become bilingual rather than monolingual in Azerbaijani. The government is already taking steps to improve teaching of Azerbaijani in community schools. But there is a place for non-governmental organizations to work with both the government and local communities to promote the heritage languages. I have been consulting with one such NGO, Üfüq-S (see ⟨http://www.ufuq.az/⟩).

Since its founding in 2004, Üfüq-S has worked with educators, leaders of cultural centers and community leaders to develop orthographies for a number of less-widely-spoken languages. It has been able to provide training in vernacular education for an Udi teacher of Russian. This teacher has now produced colorful Udi-language materials that were published by the Ministry of Education, and is using them in the Udi schools. Writers' workshops have been held in Xinaluq village where Khinalug speakers wrote and printed stories in their own language. A dictionary workshop was held in Baku for members of the Udi, Khinalug, Budukh and Lezgi communities.

Through activities like these, Üfüq-S and similar organizations can help speakers of less-widely-used languages expand the domains in which their languages are used. Activities involving literacy, which could be seen as intruding into a niche that is already filled by Azerbaijani, are actually opening new niches. Literature produced for heritage language classes and in writers' workshops generally involves texts dealing with local culture and activities. As Auger has discovered in Picard, by involving professionals and expanding the domains of the traditional languages, we can hope to help communities achieve what they are asking for: stable, active bilingualism.

7. Conclusion

The investigation of the linguistic situation in Azerbaijan presented here has implications for claims about the relationship between multilingualism and language shift. While both Russian and Azerbaijani were clearly dominant languages during the Soviet period, this did not lead to large-scale shift from less-widely-spoken languages to either. Instead, the more common scenario was for diglossia to develop where Azerbaijani was used in official situations, and the heritage language was used in the home and community. The role of Russian was restricted to urban areas even during the Soviet period, unless there were specific local conditions (as in the case of the Christian Udi) that led to its use in formal situations.

The failure of the colonial language, Russian, to play a greater role in Azerbaijan can be traced to an ecology of language during the colonial period, in particular as a development of colonization patterns. As is the case in clearly exploitation colonies, there were never enough Russian settlers to reshape the economic sphere in such a way that Russian would be needed by the average person wanting to participate in the economy.

Since independence, shift from less-widely-spoken languages to Azerbaijani has accelerated in spite of the fact that there has been no official push to force people to use Azerbaijani. This too can be explained in terms of language ecology.

In an independent Azerbaijan, knowledge of Azerbaijani has become crucial for greater involvement in the economy; the role of Azerbaijani in the ecology of language has changed. So even though the government has supported the use of heritage languages, and has spoken positively of a multi-cultural country, speakers of less-widely-used languages are shifting to Azerbaijani on their own. They see a need to know Azerbaijani, and do not see an economic need for their heritage language.

At the same time, most members of these communities express a desire to see their children become bilingual in both Azerbaijan and the traditional language. By working with professionals and other community members to find ways in which the ecological niches of the traditional languages can be expanded, it is possible that we can help the local communities see their desires fulfilled.

References

Ağacamal, Soltanov & Məmmədxan, Soltanov. 2007. *Tatca-Azərbaycanca/Azərbaycanca-Tatca lüğət*. Baku, Azerbaijan: Nurlan.

Altstadt, Audrey L. 1992. *The Azerbaijani Turks: Power and Identity under Russian Rule* [Studies of Nationalities]. Stanford CA: Hoover Institution Press.

Arutiunov, Sergei. 1998. Linguistic minorities in the Caucasus. In *Linguistic minorities in Central and Eastern Europe*, Christina Bratt Paulston & Donald Peckham (eds.), 98–114. Clevedon: Multilingual Matters.

Auger, Julie. 2011a. The impact of language revival on linguistic structure: Neuter subject pronouns in Picard. *University of Pennsylvania Working Papers in Linguistics* 17(2), Art 3. ⟨http://repository.upenn.edu/pwpl/vol17/iss2/3⟩ (21 September, 2011).

Auger, Julie. 2011b. Traditional speakers, militants, and language change. Paper presented at the Methods in Dialectology 14, University of Western Ontario.

Clifton, John M. 1997. The Kaki Ae language. In *Materials on Languages in Danger of Disappearing in the Asia-Pacific Region No. 1: Some Endangered Languages of Papua New Guinea: Kaki Ae, Musom, and Aribwatsa*, Stephen A. Wurm (ed.), 3–65 [Pacific Linguistics D-78]. Canberra: Australian National University.

Clifton, John M. 2009. The future of the Shahdagh languages. *International Journal of the Sociology of Language* 198: 33–45.

Clifton, John M.; Clifton, Deborah A.; Kirk, Peter & Ljøkjell, Roar. 2002. The sociolinguistic situation of the Udi in Azerbaijan. In *Studies in Languages of Azerbaijan*, Vol. 1, John M. Clifton (ed.),107–123. Baku, Azerbaijan and St. Petersburg, Russia: Institute of International Relations, Academy of Sciences of Azerbaijan and SIL International, North Eurasia Group. ⟨sil.org/silesr/abstract.asp?ref=2005–014⟩ (13 October 2011).

Clifton, John M.; Deckinga, Gabriela; Lucht, Laura & Tiessen, Calvin. 2003. The sociolinguistic situation of the Tat and Mountain Jews in Azerbaijan. In *Studies in Languages of Azerbaijan*, Vol. 2, John M. Clifton (ed.), 93–161. Baku, Azerbaijan and St. Petersburg, Russia: Institute of International Relations, Academy of Sciences of Azerbaijan and SIL International, North Eurasia Group. ⟨sil.org/silesr/abstract.asp?ref=2005–017⟩ (13 October 13 2011).

Clifton, John M.; Lucht, Laura; Deckinga, Gabriela; Mak, Janfer & Tiessen, Calvin. 2002. The sociolinguistic situation of the Lezgi in Azerbaijan. In *Studies in Languages of Azerbaijan*, Vol. 1, John M. Clifton (ed.), 21–46. Baku, Azerbaijan and St. Petersburg, Russia: Institute of International Relations, Academy of Sciences of Azerbaijan and SIL International, North Eurasia Group. ⟨sil.org/silesr/abstract.asp?ref=2005–012⟩ (13 October 13 2011).

Clifton, John M.; Tiessen, Calvin; Deckinga, Gabriela & Lucht, Laura. 2003. The sociolinguistic situation of the Talysh in Azerbaijan. In *Studies in Languages of Azerbaijan*, Vol. 2, John M. Clifton (ed.), 51–91. Baku, Azerbaijan and St. Petersburg, Russia: Institute of International Relations, Academy of Sciences of Azerbaijan and SIL International, North Eurasia Group. ⟨sil.org/silesr/abstract.asp?ref=2005–009⟩ (13 October 2011).

Crowley, Terry. 1995. Melanesian languages: Do they have a future? *Oceanic Linguistics* 34: 327–344.

Dixon, Robert M.W. 1991. The endangered languages of Australia, Indonesia and Oceania. In *Endangered Languages*, Robert H. Robins & Eugenius M. Uhlenbeck (eds.), 229–255. New York NY: Berg.

Fierman, William. 2005. Kazakh language and prospects for its role in Kazakh "groupness". *Ab Imperio* 2: 393–423.

Garibova, Jala. 2009. Language policy in post-Soviet Azerbaijan: Political aspects. *International Journal of the Sociology of Language* 198: 7–32.

Garibova, Jala & Asgarova, Matanat. 2009. Language policy and legislation in post-Soviet Azerbaijan. *Language Problems and Language Planning* 33: 191–217.

Grenoble, Lenore. 2003. *Language Policy in the Soviet Union* [Language Policy 3]. Dordrecht: Kluwer.

Grjunberg, A.L. & Davidova, L.X. 1982. Tatskij jazyk. In *Osnovy iranskogo jazykoznanija. Novoiranskie jazyki zapadnaja gruppa, prikaspijskie jazyki*, Vol. 3, V.S. Rastorgueva (ed.), 231–286. Moscow: Akademija Nauk SSSR.

Karan, Mark E. 2011. Understanding and forecasting Ethnolinguistic Vitality. *Journal of Multilingual and Multicultural Development* 32: 137–149.

Kosvena, M.O. (ed.). 1960. Lezginy. *Narody mira: Etnograficheskie ocherki*, Vol. 1, Narody Kavkaza, 503–519. Moscow: Izdatel'stvo Akademii Hayk SSSR.

Landau, Jacob M. & Kellner-Heinkele, Barbara. 2001. *Politics of Language in the Ex-Soviet Muslim States: Azerbayjan, Uzbekistan, Kazakhstan, Kyrgyzstan, Turmenistan and Tajikistan*. Ann Arbor MI: University of Michigan Press.

Lewis, M. Paul. (ed.). 2009. *Ethnologue: Languages of the World*, 16th edn. Dallas TX: SIL International. ⟨www.ethnologue.com⟩

Mammedov, Novruzali. 2006. *Talyshsko-Russko-Azerbajdzhanskij slovar'*. Baku, Azerbaijan: Nurlan.

Marquardt, Kyle L. 2010. Nation-building and language policy in post-Soviet Azerbaijan. *Azerbaijan in the World: ADA Biweekly* 3(19): 1–6.

Marquardt, Kyle L. 2011. Framing language policy in post-Soviet Azerbaijan: Political symbolism and interethnic harmony. *Central Asian Survey* 30: 181–196.

Moseley, Christopher (ed.). 2010. *Atlas of the World's Languages in Danger*, 3rd edn. [Memory of Peoples]. Paris: UNESCO.

Mufwene, Salikoko S. 2008. *Language Evolution: Contact, Competition and Change*. London: Continuum.

Mühlhäusler, Peter. 1987. The politics of small languages in Australia and the Pacific. *Language & Communication* 7: 1–24.

Neroznak, V.P. (ed.). 1994. *Krasnaja kniga jazykov narodov Rossii: Entsiklopedicheskij slovar'-spravochnik* (Red Book of the Languages of Russia: The Encyclopedia). Moscow: Academia.

Qübatov, Ə.B. 1991. *Azərbaycan Xalqı ilə Vahid Qardşlıq Ailəsində*. Baku, Azerbaijan: Elm.

Radloff, Carla F. 1991. *Sentence Repetition Testing for Studies of Community Bilingualism*. [Summer Institute of Linguistics and the University of Texas at Arlington Publications in Linguistics 104]. Dallas TX: Summer Institute of Linguistics and University of Texas at Arlington.

Rastorgueva, V.S. 1979. *Osnovy iranskogo jazykoznanija*, 1–4. Moscow: Akademija Nauk SSSR.

State Statistical Committee of the Republic of Azerbaijan. 2012. Population by ethnic groups, by native language and freely command of languages. ⟨www.azstat.org/statinfo/demoqraphic/en/index.shtml#⟩ (2 March 2012).

Swietochowski, Tadeusz. 1995. *Russia and Azerbaijan: A Borderland in Transition*. New York NY: Columbia University Press.

Trosterud, Trond. 1998. Folketeljingar – Väestönlaskut – Census data. ⟨http://www.hum.uit.no/a/trond/isam.html⟩ (19 October 2011).

The World Factbook. 2011. Washington DC: Central Intelligence Agency. ⟨https://www.cia.gov/library/publications/the-world-factbook/index.html⟩ (30 November 2011).

Vahtre, Lauri & Viikberg, Jüri. 1991. *The Red book of the Peoples of the Russian Empire*. ⟨http://www.eki.ee/books/redbook/⟩ (18 October 2011).

de Waal, Thomas. 2010. *The Caucasus: An Introduction*. Oxford: OUP.

Revitalizing languages through place-based language curriculum

Identity through learning*

Joana Jansen[1], Janne Underriner[1] & Roger Jacob[2]
[1]University of Oregon / [2]Wapato School District, Washington

This paper discusses the components of Identity Through Learning (ITL), language learning and curriculum development that is community centered, experiential, and collaborative. It discusses three examples of culture- and place-based curriculum projects developed at the Northwest Indian Language Institute (NILI) at the University of Oregon. We propose that place-based curriculum grounds student learning in their experiences in local events and places, and fosters community connection to traditional lifeways. As such, it can strengthen student self-esteem and identity. The paper addresses how place-based curriculum reinforces Native educators' goals for student learning, and how ITL is a promising strategy for supporting students in the classroom and beyond.

1. Introduction

Place-based curriculum provides a meaningful educational experience for Native students as it promotes authentic learning that supports communities in revitalizing their world views and associated lifeways. It honors the connection to one's home, family, community and world. The traditions embodied in the curriculum provide confidence and grounding for the child or adult learner while providing a perspective from which to investigate and understand the world at large.

* We would like to thank all the language teachers and students who have applied this work in their teaching and have made suggestions for its improvement. We are honored to be a part of efforts to promote learning, foster curiosity and develop a connection to community. We additionally thank the funders who have supported our various curriculum projects; they are named in conjunction with the projects. This chapter is based in part on work supported by the National Science Foundation under Grant Number 1064459. We also warmly acknowledge Judith Fernandes and Zalmai Zahir for their contributions to this paper. We thank the editors of this volume for their thoughtful comments.

Place-based language learning is an essential tool for language revitalization as it fosters community connection to traditional lifeways. With community at the center, students learn about core values, culture, ancestral and home lands, and their people's history as they learn their language. Students become connected to what is essential to their tribal community and to the ways of their ancestors. Place-based learning is experiential and so nurtures students' curiosity, builds cooperation among students, and strengthens problem solving abilities. It links students with members of their community who contribute to its diverseness, and in so doing it opens students' awareness to elders, leaders, and mentors they might not have encountered in a more teacher-centered classroom learning experience. Culture and place-based learning is more compatible with the way information was taught or passed down in the Pacific Northwest of the United States of America 150 years ago, before the time of federally imposed education. It engages youth and children in learning their language in culturally appropriate ways that are participatory and project based. It builds relationships among mentors and youth, and older children as role models for younger children.

This paper introduces readers to the concepts of place-based curriculum and addresses how place-based curriculum supports Native educators' goals for their students. We describe the context of our work in the Pacific Northwest of the United States, and discuss the components of what we refer to as *Identity Through Learning (ITL)*. The paper then provides three examples of place-based curriculum projects developed at the Northwest Indian Language Institute (NILI) at the University of Oregon, and we address ways that linguists and documentation specialists support curriculum development. The paper concludes with a discussion of Identity Through Learning as ITL is a promising strategy for supporting students in the classroom and beyond.

We begin with a brief introduction to the Native languages situation in the Northwest, Oregon and Washington in particular, touching historically on languages before contact with European-Americans and ending with the current status of these languages and the circumstances in which they are taught.

2. Native languages in the Pacific Northwest

2.1 Loss and revitalization

In 1800, before contact with settlers, what is now Oregon and Washington were inhabited solely by people currently referred to as Northwest Coast, Plateau, and Great Basin Native Americans. Historically, the entire Pacific Northwest region had an incredibly diverse set of languages with some 25 plus languages being spoken in Oregon alone, many with multiple dialects (Hymes 2007). Of these, only

a handful of languages with living speakers remains. What is now Washington State contained seven language families with around 23 languages and multiple dialects (Kinkade et al. 1998; Thompson & Kinkade 1990). The last fluent elders of many Oregon languages, including Alsea, Yaquina, Tualatin, Yonkala, Tillamook, Chinook, Kathlamet, Clatskanie, Rogue River, Molalla, Takelma and Cayuse, died during the first half of the twentieth century. The last fluent speaker of Hanis and Miluk Coos passed in the early 1960s and the last fluent elder speaker of Klamath in the early 2000s. Today, the only speakers of Kiksht and Walla Walla in Oregon are second language learners. Each of these languages represents a unique view of the world – the key and summation of an entire culture's long history.

After the arrival of White trappers, traders, military, and settlers, disease left some languages with very few remaining speakers. The reservation system further destabilized languages and communities by placing people of disparate cultural and linguistic backgrounds together. For example, on the Siletz and Grand Ronde reservations there were people from over a dozen different tribes, speaking over a dozen mutually unintelligible languages. In this situation, Chinuk Wawa was spoken by most people, and became the community language; the original languages of those who were forced to the reservation fell out of daily use. Government educational policies added to the loss, as children were taken from their homes, often forcibly, and sent to boarding schools where the use of their native language resulted in harsh punishment. Children were forced to learn and speak only English. When these children became parents themselves, they wanted to spare their children the pain they had experienced, and so did not encourage fluency in the native language.

Tribal communities are acutely aware of their loss. Communities that have not had living speakers for generations are trying to recover whatever they can of their languages from minimal written and audio records. The languages with fluent speakers are acting with urgency to revitalize their languages. Communities are faced with the urgent task of ensuring the survival of their languages for future generations. With current efforts communities are once again hearing their languages spoken. Children and adults both are speaking Ichishkíin (Sahaptin), Coastal and Inland Salish, Numu (Northern Paiute), Chinuk Wawa, Siletz and Tolowa Dene (Athabaskan), Kiksht (Wasco), Klamath and Nimiipuu (Nez Perce). Children are finding their identity in learning from their elders.

2.2 Native language in the classroom

Native American students fall behind other ethnic groups in math and reading attainment, high school graduation and college entry rates (NEA 2011). Most Native children live in homes where their language is not known or used and attend schools where their language, cultural traditions and values are undermined. Loss

of cultural identity, and teaching practices that do not resonate with their culture have been linked to the lack of success Native American children are experiencing in education (Wiley 2008).

Current studies indicate how integral language is to the sense of well-being of Native children, and in turn, to their academic performance, self-esteem, and ability to navigate in a complex world. Studies also show that connecting Indian youth with their languages increases their resiliency to addiction and promotes positive health and well-being (see for example Goodkind et al. 2011). According to Mmari, Blum & Teufel-Shone (2010), risk factors for youth include loss of language and culture, while protective factors include knowing one's Native language, participating in traditional ceremonies, and dancing at powwows.

Place-based education can therefore boost student achievement and well-being. The examples and discussion of curriculum that follow have been used in formal learning situations: schools, preschools and small group lessons. These formal settings are not the only context for language learning or place-based models. Furthermore, they have not typically been successful on their own in bringing the languages back to everyday use throughout the community. Without language spoken in homes and throughout communities, revitalization successes are limited. Hinton sums this up: " the most important locus of language revitalization is not in the schools, but rather the home, the last bastion from which language was lost, and the primary place where first language acquisition occurs" (Hinton 2013: xiv).

However, school-based revitalization efforts are critical for several reasons. First of all, in the communities in which we work it is the only place that many students have access to their language and traditions. As the Native languages in the Northwest United States are severely endangered, the majority no longer spoken since the mid-1900s, learning them through intergenerational transmission in the home and family is no longer viable. It is out of necessity that Native languages have found their way into public and tribal schools; it is this unique, or perhaps one could argue, unnatural learning situation that drives communities to develop meaningful and culturally rooted teaching and learning strategies for Native languages. For example, at the Confederated Tribes of Grand Ronde, the immersion programs have made it possible for many more students to learn the language and culture in an effective learning environment than if language transmission was restricted to the home only. If there were no classroom teaching, most of the children in the program would have no Native language learning. A second important strength of formal learning situations is that they can identify and inspire those community members who are dedicated to bringing the language home. Finally, having language classes in schools and at language programs can boost awareness and support for the language throughout the community, leading to increased opportunities for revitalization.

3. Models for Northwest language learning

3.1 Place-based curriculum

A place-based educational approach grounds curriculum and lessons in students' experiences in local events and places, and acknowledges that learning happens not only in formal educational settings but also outside of school in families and communities. This reinforces connections to one's home, family, community and world. Included components can be the cultural, historical, social, religious and/or economic relevance of specific locations or areas (Smith 2002; Gruenewald 2003).

Gruenewald writes that place-based education does not have one particular theoretical tradition, but rather that

> its practices and purposes can be connected to experiential learning, contextual learning, problem-based learning, constructivism, outdoor education, indigenous education, environmental and ecological education, bioregional education, democratic education, multicultural education, community-based education, critical pedagogy itself, as well as other approaches that are concerned with context and the value of learning from and nurturing specific places, communities, or regions (2003: 3).

Place-based learning addresses a long-stated criticism of most educational experiences: that outside of school, people "experience the world directly; in school, that experience is mediated, and the job of students…is to internalize and master knowledge created by others" (Smith 2002: 586). A place-based curriculum is intimately dependent on the world outside the classroom and is responsive to its locality. Place-based education has a goal of involving "teachers and students in the firsthand experience of local life and in the political process of understanding and shaping what happens there" (Gruenewald 2003: 620).

Place-based education has links to communicative and culturally-based approaches. Communicative approaches to language teaching (see Brown 2006; specifically for Native language teaching see Hinton & Hale 2001; Supahan & Supahan 2001) stress the significance of authentic communication. This can easily be imbedded in a place-based curriculum.

Place-based education supports recommendations of Native educators for Native students. The National Education Association proposes that Native ways of knowing be incorporated as a "critical cornerstone of relevant, rigorous, and high quality instruction for Native students" (2010: 4). Place-based education meets the call for integration of the local and the inclusion of cultural knowledge in teaching, as well as increased involvement by the community (Blanchard 1999; Gay 2000; Nee-Benham & Cooper 2000). The traditional importance of place is discussed by Cajete (1994), who writes that the purpose of traditional education in Native cultures is to deeply connect young people to their heritage and their physical

homelands. Curriculum geared toward exploring places can deepen empathetic connections and expand the possibilities for learning outward. Sobel explains, "[place-based] curriculum can mirror the expanding scope of the child's [or adult's] significant world, focusing first on the home and school, then the neighborhood, the community, the region, and beyond" (1996: 19).

Clearly, the literature points to a curricular design that considers both culture and place to ensure a meaningful educational experience for Native students. Demmert and Towner offer an operational definition of culturally-based education that brings together culture and place. It includes six critical elements (2003: 9):

1. Recognition and use of Native American languages.
2. Pedagogy that stresses traditional cultural characteristics, including the importance of adult-child interactions.
3. Teaching strategies that are based on traditional culture as well as contemporary ways of knowing and learning.
4. Curriculum based on traditional culture that recognizes the importance of Native spirituality, and places the education of young children in a contemporary context (e.g. use and understanding of the visual arts, legends, oral histories, and fundamental beliefs of the community).
5. Strong Native community participation in educating children and in the planning and operation of school activities.
6. Knowledge and use of the social and political mores of the community.

Perhaps of most importance, tribal elders highlight the notion that place-based education brings about wholeness of the individual and community, and thus contributes to students' positive self-esteem and identity. Elders note that younger people do not all know the cultural part of life; who they are; how their ancestors survived. Teaching to the seasonal calendar links what was and is traditionally done during each season to distinct geographical areas. Other native educators with whom we work design curriculum centered around traditional foods and nutrition, longhouse and sweathouse protocol, and legends that link powerful moral lessons with sites on traditional lands.

3.2 Identity through learning

Place-based learning is not a new trend within education or Native education, but its emergence in Northwest Native communities as a promising approach for language teaching is a more recent development. An example of a Northwest place-based curriculum, albeit one that does not teach language, is the Salmon Watch curriculum prepared by Oregon Trout, an organization with the mission of protecting and restoring native fish and ecosystems. The organization is a

public-private partnership with input from the Columbia River Inter-Tribal Fish Commission and individual tribal members. The middle school through high school curriculum was "designed to provide a holistic, multi-disciplinary and watershed-based approach to environmental education, using the salmon as the key indicator species of watershed health and the cultural icon of the Pacific Northwest" (Oregon Trout 2005: iii). The school year begins with planning and taking a field trip to a riparian environment. While there, students may make observations, collect data, or speak with a tribal member about the significance of salmon or the selected site. During the year, students learn about salmon lifecycle, Native American storytelling, watersheds, and human effects. Students also carry out a service learning project and disseminate the results.

Looking farther north, the Alaska Native Knowledge Network (ANKN) at the University of Alaska Fairbanks has been a forerunner in promoting Alaska Native ways of knowing and provides resources on integrating Native and Western knowledge systems within educational systems. The materials and school standards the ANKN has developed emphasize a shift from learning *about* cultural heritage to learning *through* the local culture as a foundation for all education (Barnhardt 2005).

The Northwest Indian Language Institute (NILI) at the University of Oregon (UO) began writing curriculum with tribes in 1997 with a focus on topics directly related to culture and culture revitalization. The Institute's ongoing collaborative efforts support and strengthen language preservation in various communities. Projects range from linguistic documentation and revitalization to curriculum and standards development to language program and state policy development.

By 2000, we realized that NILI and partners were addressing not only culture but also its link to geographical areas and to individuals in the community. This motivated us to become more aware of the literature on place-based learning, and with the Confederated Tribes of Grand Ronde (CTGR), we began developing culture and place-based curriculum for their Chinuk Wawa immersion pre-school. Over the past 12 years, with funding from the Administration for Native Americans (ANA), Department of Education, and Spirit Mountain Community Fund (SMCF), NILI and the CTGR Cultural Resources and Education departments have developed and implemented culture and place-based curriculum which focuses on language arts, ethnoscience, social studies/history and math skills development.

Through SMCF funding in particular, NILI, in collaboration with communities, has been able to develop an Identity Through Learning (ITL) framework. It is a place-based learning model that is rooted in the lifeways of indigenous communities. We have identified three elements that are integral to the ITL

framework and essential for curriculum design and development. For us, ITL is *Community-centered, Experiential,* and *Collaborative*. We address each of these three in the descriptions of curriculum projects below.

Identity Through Learning curriculum is *Community-centered.* We assume a broad definition of place to include home, family, community, and land (the area within current reservations as well as traditional lands). The examples below show that a seasonally based curriculum links learners to individuals and places locally on the reservation and more broadly to ancestral or ceded lands. Specific locations throughout what is now Oregon and Washington have a rich and sustained history as locations for gathering and celebrations, religious practices, fishing, hunting, and food collecting. Many geographical formations along the Columbia River, for example, are tied to a legend that describes how they came into being or an event that took place there (Aguilar 2005). A drive with an elder along the river brings out these legends, as well as stories of events that took place within the elder's lifetime: childhood fishing sites; the loss of those sites when the river was dammed; locations of petroglyphs; more recent events and activities with family and friends. We also take community-centered to mean that the curriculum represents what community members believe is important for their children to know. The multidisciplinary character of community-centered curriculum means it relies on the input of a wide section of community experts. The units described below are not limited to language, but involve art, math, botany, physics, and more.

Identity Through Learning is *Experiential*. It takes students out of the classroom and places them in their environment. It asks them to engage and create. For Kolb and Fry (1975) experiential learning follows a spiral. The spiral begins with experience, then moves through reflection, forming and testing concepts, and reaching conclusions that are applied to the next iteration of the cycle. In the canoe curriculum discussed below, learners learn how to move around properly in the canoe: how to get in and out, how to move from one side to the other, how to bail. As children experience the effect their weight and movement have on the boat's movement they are able to adjust their actions to keep the canoe balanced.

Experiential learning requires the student to be inquisitive and interested in life, their learning and their larger environment. In the basketry project, for example, students gather and prepare materials and in so doing gain respect for the materials and an understanding of how they relate to the natural environment. The curriculum we develop is typically intergenerational, and this too is experiential, as students work with elders and hear their words and experiences. This helps youth internalize a worldview rooted in their ancestral legacy. The curriculum is hands-on, and we have learned that hands-on activities can both energize and calm students as they work. Hands-on activities thus contribute to classroom management. This way of learning inspires sharing, conversation,

curiosity, and group work at all age levels; with younger children, a child can sit on an elder's lap to weave or hold an elder's hand when exploring their environment which provides feelings of security or nurturing.

Thirdly, Identity Through Learning is *Collaborative*. Collaborative work is a core value of NILI and the need for and benefits of this sort of work are increasingly apparent across academic fields. Grinevald (2003: 57) notes the evolution towards linguistic fieldwork done *with* and *by* language community members, rather than *on* a language or *for* the language community. Cameron et al. (1992) and Rice (2006) discuss an "empowerment" framework for fieldwork, in which "the work is on the language, for the speakers, and *with* the speakers, taking into account the knowledge that the speakers bring and their goals and aspirations in the work" (Rice 2006: 132). Yamada (2010) describes a Community Partnerships Model, a collaborative approach that depends on long-term partnerships in which projects are mutually determined and mutually beneficial. The desired result is a partnership between researchers and speech communities in jointly planned and jointly beneficial projects.

In our work, the collaborative nature of effective curriculum design means that it necessitates a team of, for example, language program staff, elders, topic specialists, linguists, curriculum writers, school district representatives, teachers, parents and students. Cooperation and communication are essential throughout the process, and team members respect each other's views and contributions to the team. So, while NILI staff may be seen as the outside academic curriculum and documentation "experts", this is no more important than the knowledge and expertise the speech community partners bring, whether that is about the language to include, specific processes related to the topic, or the best way to engage middle-school students. We note as well that these divisions between groups are not clear cut; one of this paper's authors and many of the curriculum developers named below are both academic and speech community experts, and we increasingly work with, train, and learn from people who are members of both academic institutions and speech communities.

The ITL curriculum we develop is typically based on materials that have been collected with the goals of documenting and describing a language as well as revitalizing and teaching it. We discuss the use documented materials in curriculum development for each project below. Language documentation and language teaching go hand in hand in our work, and by acknowledging that at the project outset, we end up with rich materials that support multiple users (see also Jansen & Beavert 2010; Yamada 2011).

Before we turn to look at three examples of how place-based curriculum is contributing to language revitalization, we address teaching contexts in the Northwest.

3.3 Oregon and Washington teaching contexts

Linguists and curriculum developers need to know what kinds of programs are in place or desired in the area when assisting to develop classroom materials or deciding what sorts of documentary materials can support teaching. Some teachers and classrooms adhere closely to a second language communicative approach based on comprehensible input. Their goal is day-to-day communication in the target language. In some situations, the goal of teaching language and culture is to strengthen self-esteem and provide a heightened awareness of culture, place, and history. Teaching may focus more on learning vocabulary and phrases for situations of deep cultural relevance – for example, words and phrases that are used in religious ceremonies or while gathering food. The desire and ability to use technological aids for language learning also varies by teacher and available technology.

Many Northwest teachers are working to incorporate immersion teaching into their classrooms. A traditional model of immersion in which students meet for a day or half-day and everything in the classroom happens in the target language is not realistic for all Northwest language situations: presenting rich content and culture requires a higher degree of fluency than many teachers have. In classrooms on the Yakama Nation, for example, language teachers are typically language learners, younger adults who have a strong commitment to their language and the energy to teach it. Their challenge is to keep at least a step ahead of their students, providing a language-rich classroom environment given their own level of proficiency.

The benefits of using immersion techniques for a shorter time are available to less than fully fluent teachers. In these situations immersion teaching calls for a strategy of beginning with using the target language perhaps 15 minutes at a time and increasing from there. Hinton suggests that a teacher who is learning her own language while she is teaching it focus on learning various components of a lesson. If a teacher learns the lesson elements – not only the new and review material presented in the lesson but also greetings, classroom management vocabulary, and informal patter – she can have an immersion classroom (Hinton 2003: 80). Another technique uses specific activities to stretch what teachers do know. Zalmai Zahir, in a Lushootseed immersion and methods class taught at NILI's Summer Institute, demonstrated for teachers how nothing but counting from 1–10 could be a ten minute activity that maintained student interest throughout with song, humor and physical movements.

Oregon and Washington State Departments of Education acknowledge that Native languages are under the scope of tribal governments and offer special certification for Native language teachers. In both states, the tribes determine who should be credentialed to teach language in public schools. Teachers must meet

a tribally determined level of proficiency in the language, but are not required to complete a university-level teacher training program. These certification programs allow for Native languages to be taught in public schools and give tribes the authority over their own languages. Elders without formal schooling can teach in the classroom and younger teachers, who are often not able to leave their communities for a four to five year course of study, can maximize precious and limited time with their elders. However, a drawback is that teachers with a native language teacher license may lack teacher training and ongoing professional development opportunities.

Teaching and learning situations in the Northwest range from formal classroom settings to informal in-home language learning. Some teachers see their students five days a week, others only for a half-hour a week. Although what is presented in this paper is based on teaching language in classroom situations, this is not the only context for language revitalization. The material is adaptable to other learning situations, such as home and Master-Apprentice settings and community programs. Curriculum developed using the Identity Through Learning framework supports all of these teaching contexts.

4. Curriculum examples: Basketry, plants, canoes

This section discusses three place-based language-teaching units that represent the Identity Through Learning framework. For each, we give an overview of the project, then discuss its goals and objectives and the resulting products.

4.1 Grand Ronde Basketry: Place, community and voices – intergenerational learning

Overview:

The Northwest Indian Language Institute (NILI) was awarded a grant from Spirit Mountain Community Fund (SMCF) in 2008 to support the development of basketry curriculum for the Confederated Tribes of Grand Ronde (CTGR) and other tribes of the Northwest. *Basketry: Place, Community and Voices* is a multidisciplinary, year-long unit on basketry of the Grand Ronde people. It met the need for place-based curriculum at CTGR's Preschool, Kindergarten, Chinuk Wawa, After-School and Title VII programs as well as the Tribe's Library by providing curriculum and instruction that focuses on the resources and people of the Grand Ronde community.

As mentioned earlier, NILI began working with CTGR on culture and place-based curriculum in 2000 as a means to revitalize the world views and associated

lifeways of the Tribe. What we learned from our earlier projects was that more participation from tribal elders and community members was needed at the curriculum planning stage. So for this project staff from the Cultural Resources Department, the Chinuk Wawa program, the After School program, and the Tribal Library as well as community elders and basket weavers participated in planning, developing and teaching the curriculum. The intergenerational relationships fostered by the project provided opportunities for elders to share their knowledge of culture and place directly with students. The project team consisted of: Connie Graves, Dolores Parmenter, Tony Johnson, Kathy Cole, Bobby Mercier, Marion Mercier, Jeanne Johnson, Judith Fernandes, Janne Underriner, Henry Zenk and Margaret Matthewson.

Resources collected in earlier linguistic and anthropological projects were important to the project. Some of the material Dr. Henry Zenk, the Tribe's linguist, recorded in 1981 with elders included information about gathering areas, material/hazel processing for weaving, using baskets in gathering, and tools for basket making. We relied on language documented by Tony Johnson, then the Tribe's Cultural Education Coordinator and Zenk for development of the curriculum. They, as well as the teachers, transcribed the curriculum and stories in Chinuk Wawa.

Goals and objectives:

The project goal that emerged from our community meetings was to immerse Grand Ronde community children (and others) in a learning environment that centered on hazel and juncus basketry and that reflected the language and values of the Tribe. Curriculum was taught at the Tribes' Preschool, Kindergarten, After-School and Title VII programs as well as at two workshops that focused on family learning. The main objectives of the project were for students to understand that baskets are an important part of Grand Ronde culture; to know that juncus and hazel are used in creating traditional baskets; to identify different weaving materials in situ and in class, and be able to name them in both Chinuk Wawa and English; to explain in Chinuk Wawa various stages of the weaving process.

The curriculum met Oregon State standards in math, science, social studies, history, art, and literary arts. Dependent on the extent of prior language and cultural experience, we expected that each student would increase skills (proficiency) in the following ways: (1) develop better small motor control; (2) gain knowledge of (or be able to identify) (a) where on the reservation to collect materials, (b) how to process cedar; (c) the stages of weaving; (d) design elements; (e) functionality; (f) cultural role of cedar/basketry in general; (g) the distinction a basket weaver holds in the Grand Ronde community; (h) the cultural and historical role of an artist in the Grand Ronde community.

Judith Fernandes, NILI's head teacher trainer and curriculum developer, worked with the team to establish the curriculum's targeted learning skills in the following areas:

Math – Counting weavers; Counting sticks needed for a weaving project; Understanding concept of "every other one"; Geometric basket designs; Estimation; Even and odd numbers.

Social Studies and History – Use of baskets in earlier times; Current use of baskets; Basket weavers past and present; Influence of outside communities.

Stories and Literature – Pictorial biography of Hattie Hudson, a basket weaver. We go gather, a story about giving back to nature when taking from it; Basket woman (traditional story).

Art and Music – Symmetry; Form and function; Traditional design; Present day design; Gathering Song.

Science – Where, when and how to harvest; Charring sticks for bark removal; Best management practices for guaranteeing future harvests; Leaching; Boiling; Dyeing; Processing materials; Qualities of good basketry materials; Experimenting with materials.

Products:

The products that resulted from the project were a multi-disciplinary, year-long unit on basketry of the Grand Ronde people that consisted of twenty lessons. The unit was piloted and then revised. Six story books on gathering and basketry were written and illustrated by community members.

Four step-by-step pictorial books on juncus and hazel location, gathering, processing, and weaving were created in both Chinuk Wawa and English. Also, two community weaving workshops were video recorded. A portable educational box containing all materials from the project was assembled. All materials are housed at the Tribe's Library, Cultural Resources Department and at NILI. (NILI's website at pages.uoregon.edu/nwili/resources/curriculum shows additional curriculum examples.)

4.2 Tamaníksh: Yakama Nation Natural Resources catalog and curriculum

Overview:

A curriculum project underway for the Yakama language (a dialect of Ichishkíin or Sahaptin, here referred to as Ichishkíin) revolves around *Tamaníksh*, plants for food and medicinal uses. The place-based curriculum development is a component of an ongoing interdisciplinary effort to support resource management, language documentation, and language revitalization. The project began in 2005 with a grant to the Yakama Nation from the United States Forest Service to investigate ways to assess and improve forest cultural resource management. The project

included gathering input from two groups of people with Yakama forest management experience and cultural resource knowledge: Yakama elders with extensive forest cultural resource utilization experience and Yakama Indian and non-Indian resource managers employed by the Yakama Nation or Bureau of Indian Affairs. A common concern shared among Yakama elders and resource managers was a lack of identification and management of culturally significant plants and areas. It was also noted that elders learned about resources in the Ichishkíin language, and therefore the Ichishkíin language holds information and recollections in a way that the English language does not. Most of the elders responded to survey questions in Ichishkíin. The 2009 final report (Jacob 2009) called for the development of a Yakama Nation Resource Catalog in Ichishkíin and English. This catalog simultaneously strengthens natural and cultural resource management and supports teaching and preserving Ichishkíin.

With additional funding from the National Science Foundation and the Substance Abuse and Mental Health Services Administration, the catalog is now being expanded and teaching materials are being developed. Partners include the Yakama Nation Division of Natural Resources, the Yakama Nation Language Program, the Yakama Reservation Wellness Coalition, two high schools on the Yakama Reservation, and NILI. Team members are elders Átway Tayúyapam (Hazel Watlamet), Tux̲ámshish (Virginia Beavert) and Kusúmwhy (Levina Wilkins), along with Michelle Jacob, Roger Jacob, Joana Jansen, Rose Miller, Greg Sutterlict, Janne Underriner and Zelda Winnier. Classroom students at the University of Oregon have developed catalog materials; students and volunteers in tribal classrooms and Yakama Nation language program apprentices have been involved.

The materials used in the curriculum and collected for the resource catalog meet the needs of both language education and language documentation. On the classroom side, developed activities and curriculum richly support student learning. On the documentation side, materials collected contribute to the goal of having "a representative and lasting multipurpose record of a language" (Himmelman 2006: 1).

Goals and objectives:

The curriculum addressed in this section is intended for high school language students, 14–19 years of age. They are first-year students of Ichishkíin at one of several high schools on the Yakama Reservation, at Level One (beginner) proficiency based on the Northwest Indian Language Benchmarks (NILI 2008). Most are Yakama Nation members or of Yakama descent; some are non-Indian or are affiliated with other Northwest tribes. The ten-week unit (some already developed and some in planning stages) is intended for a classroom language learning setting, although we anticipate that the lessons will be useful to teachers and learners in other contexts as well. The curriculum includes elder visits

to the classroom, audio and video materials, field trips, and family interviews. Hands-on activities include gathering, processing, and preparing plant materials at the proper seasonal times. Students will also produce materials for the Yakama Nation Natural Resources catalog, adding to the database of described resources for cultural resource managers and future language learners. The unit also supports subsequent units that focus on Longhouse and Sweathouse. As students progress in their language ability as well as their knowledge of the resources they describe, they will revisit and expand catalog entries.

For the students, objectives include: recognize culturally significant plants; recognize significance of plants to the Yakama people; say 2–3 Ichishkíin sentences about a relative's practices around root foods; describe environments of 5–6 plants; name places where these plants grow; read, write, understand and say plant gathering and plant processing/preparation vocabulary and sentences; construct a catalog entry of a culturally relevant plant with picture(s), recording of an elder, and sentences.

Products:
The developed curriculum involves audio, video, and visual materials. It also includes lesson plans for each day of the unit. Language teachers are often also language learners, and so vocabulary related to classroom routines is part of the lesson plan. Teachers also relate that they are typically pressed for time, so having a prepared lesson plan with daily goals and scheduled activities assists them in adapting the materials to their classroom and learners. In addition, the inclusion of the lesson plans supports early-career teachers and language teachers without formal training. The output from students and classrooms further supports and expands the catalog of natural resources, providing materials for other students as well as resource managers.

We see that this type of curriculum provides a connection to students' culture in a deeper way: they may already go to the Longhouse or ceremonies where they see traditional foods, for example, but they only know what they look like in that one context. They may not be comfortable asking elders for more information when in a traditional setting like the Longhouse. This curriculum provides students an opportunity to learn more about traditional foods, where they are located, how to prepare and eat them while learning and speaking Ichishkíin and becoming more connected to their environment.

4.3 Lushootseed canoe curriculum

A Canoe Curriculum for the Lushootseed language and culture (Zahir 2007a, b) is a third example of a place-based curriculum that reflects Identity Through Learning. Lushootseed is spoken in the Puget Sound area of Washington State, through

the Puget Sound drainage area north to the Skagit Valley. The curriculum gives language students a brief introduction to the canoe culture of Puget Salish First People. The curriculum is designed for learners in preschool to age 8. It is a full immersion curriculum; learners may be at any stage of acquisition, but teachers need to be of intermediate-high language proficiency. It was funded by the First Nations Development Institute through its Eagle Staff Program, as well as the Potlatch Fund.

The curriculum is supported by, and supports, the Canoe Journey. Canoe Journey is a yearly event that honors the importance of the canoe in the culture for its use in trade, diplomacy, and social gatherings, and "serves to preserve traditional Northwest Native life-ways and traditions, and facilitates revitalization of language and culture" (Viles 2010: 1). It is also a vibrant example of revitalization. The use of canoes in the region faded with the arrival of white settlers and industrialization. The Canoe Journey, which began in 1989 with only a handful of canoes and tribes, has grown to an anticipated and highly regarded yearly event with up to 90 tribal groups participating and more than 100 canoes. Over the course of 2–4 weeks, canoes paddle along different Puget Sound routes, making stops at reservations and villages along the way, to the final host community. This time offers opportunities for Native language learning and use. (For more information on Canoe Journey as an event that supports language revitalization, see Viles 2010; Zahir 2007a, 2010.)

Goals and objectives:
The curriculum addresses canoe history; canoe styles and uses; canoe implements and uses; training for paddlers; canoe etiquette; speeches; brief instructions for making canoes. It also teaches cultural values: the importance of respect, strength, prayer, wisdom, courage and humility. Information on the plants, animals, and marine life seen on the canoe journey is also included.

For each unit, student objectives for vocabulary mastery, reading, writing, math, science, song and dance, and gross and fine motor skills are given. The cognitive skills and cultural values and practices that are the focus of the lesson are also laid out. For example, in one unit, students learn that there are different types of canoes in the region, and how these types are different (cognitive skills). They learn the names of the canoe categories and how and where the canoes are used; the concept of 'place' in relation to water and how the words describing place are related to water; to recognize places on land from the perspective of being on the water (cultural values). They make paper models of canoes and manipulate traditional toy canoes, and practice entering and exiting a canoe safely and respectfully (cultural practices).

Products:
The materials consist of workbooks for educators (titled Journey Curriculum and School Curriculum, available at http://www.pugetsalish.com/downloads.aspx) as well as associated hands-on materials, some that the teachers and students create, others that are objects used and seen on the canoe journey The website also includes transcribed language documentation and photos that support the lessons. This includes canoe terminology originally published in the 1920's.

5. Discussion

Each example discussed above incorporates the three elements we see as core to learners developing Identity Through Learning and experiencing their culture, language and land. Basketry: Place, Community and Voices is *community-centered* as it focuses on the resources and people of the Grand Ronde community and it serves all CTGR members, Chinuk Wawa speakers and other learners of Chinuk Wawa by making the curriculum available online. It is *collaborative* as the project staff from various departments as well as community elders and basket weavers participated in planning, developing and teaching the curriculum. Students engage in *experiential* learning as the relationships fostered by the project provide opportunities for elders to share their knowledge of culture and place directly with students. It promotes students experimenting with basketry materials and weaving techniques and experiencing first-hand locating and gathering basketry materials.

The Ichishkíin *Tamaníksh* curriculum is *community-centered,* in that its importance and expertise are rooted in the community. It is inherently multidisciplinary. Students engage in *experiential* learning, with hands-on components that lead them to come to conclusions about, for example, the way to present their plant entries or the way to prepare a root for storage, then later to revisit the topic. It is *collaborative*, with a broad range of participants, and it supports a number of community goals.

The canoe curriculum is particularly *experiential*. This curriculum stresses that learners will benefit from dynamic teaching: "Have a plan for the day, but if all the children want to do is play with the bailer and the water in the canoe, let them. They're engaged. Interact with them in Lushootseed…. Maybe the canoe rocks. Maybe bailing water makes "rivers" in the ground. Go with it and have fun. Make the experience the curriculum" (Zahir 2007a: 2). It was written with input from elders and community members and is therefore *collaborative*. It is *community-centered* as it is useable by all Lushootseed speaking people. Additionally, it merges to a

more global level with its presence at Canoe Journey, and provides inspiration for speakers of other Native languages to adopt the same model. It is a model of immersion language learning in a particular community setting.

The above curriculum examples illustrate how collaborating with communities on curriculum development yields a richer, more meaningful product. Linguists and documentation experts have much to add to these efforts. Collaboration at the onset of the project clarifies the language community's preferred teaching methods and types of curriculum design, and specifies what types of materials are useful in the classroom, given the language teaching strategies used in the community. It identifies topics that are important to the community, so that thematic units can be the rich basis for curriculum and language documentation. Further, supporting and contributing to curriculum efforts builds relationships. If community outsiders work on language documentation within a framework of respect and support for community goals, and one of these goals is classroom teaching, it is essential to support it.

Although quantitative research of Native student success is limited, ITL is a promising strategy for supporting students in the classroom and beyond. Native students have increased academic achievements when schools and classrooms validate and incorporate their culture (Demmert & Towner 2003; Lipka 2002). A recent report by the National Education Association (NEA) describes successful strategies used by Native educators to improve opportunities and achievements of Native high school students. It addresses the deep sense of place held by many Native high school students, and the importance of teachers understanding the particular context in which they teach. The report also states that an educational approach that "infuses the history, values, and language – ways of knowing- of Native people into the contents of the curriculum, the language of instruction, the delivery of instruction, and the interaction with Native students" engages students and has proven successful at boosting student achievement and retention (2011:23). A focus on Native culture can boost family participation and attendance (NEA 2010).

NILI staff are developing measures to evaluate the effectiveness of culture and place-based curriculum, and language proficiency assessments to measure oral and literacy skills. At the Confederated Tribes of Grand Ronde's K-1 Chinuk Wawa immersion program, initial language assessments are given in the fall and then again in the spring; these address language proficiency. Each place-based unit has pre and post tests to measure content and concept retention. They are limited to five questions developed by the teachers; for example, "Why do beavers have big teeth? Where do beavers live?" Students' answers are grounded in real life experience as they visit areas on the river where beaver dams are located and they examine beaver pelts (which are part of their classroom environment).

Overall, Grand Ronde K-1 teachers prefer informal assessment, that of observing the progress or skill development of their students, as they find it provides more meaningful information about their delivery of the curriculum and unit activities, student grasp of the concepts and student learning styles. Additionally, teachers keep portfolios on each student for documentation of growth and progress. Teachers report that their students are enthusiastic about their learning because it validates the environment and their lives.

Our informal experiences suggest that place-based learning increases community interest and contributes to learners' wellbeing. A continuing interest of the authors is to better understand the relationship between language revitalization and the overall health, including identity and self-esteem of learners. Over our years of collaboration, NILI and tribal programs have witnessed examples of how increasing one's knowledge of language and culture is related to an increase in self-esteem and cultural pride. Elders acknowledge that this association is evidence of the importance of language learning.

In an effort to begin evaluating how place-based education promotes self-esteem, NILI, high school Ichishkíin language teachers, Yakama elders and the Yakama Nation's Wellness Coalition have developed a study to measure whether education rich in language and culture prevents (acts as a protective factor against) drug and alcohol abuse. The study was recognized by the Native American Center for Excellence in 2011, and the pilot at the Yakama Tribal and E.A.G.L.E. high schools in Toppenish, Washington was funded by the Substance Abuse and Mental Health Services Administration in 2012. A distinguishing factor of our project is that it reflects the values of Identity Through Learning – it is collaborative, experiential, community centered.

These are on-going studies and therefore our findings are pending. It is our hope that the approach outlined in this paper contributes to communities' efforts to revitalize languages and culture, and more importantly, to foster healthy youth through a connection to traditional lifeways.

References

Aguilar, George. 2005. *When the River Ran Wild!: Indian Traditions on the Mid-Columbia and the Warm Springs Reservation.* Portland OR: Oregon Historical Society Press.

Cajete, Gregory. 2000. *Look to the Mountain: An Ecology of Indigenous Education*. Durango CO: Kivakí Press.

Cameron, Deborah; Frazer, Elizabeth; Harvey, Penelope; Rampton, M. B. H.; & Richardson, Kay. 1992. *Researching Language: Issues of Power and Method*. New York: Routledge.

Barnhardt, Ray. 2005. *Creating a Place for Indigenous Knowledge in Education: The Alaska Native Knowledge Network*. Fairbanks AK: ANKN.

Blanchard, Rosemary Ann. 1999. Forging a civic relationship between Native youth and their Indigenous Nation: A tribally-specific, tribally-formed social studies curriculum. In *Indigenous Education Around the World, Workshop Papers from the World Indigenous People's Conference,* Ray Barnhardt (ed.). Fairbanks: University of Alaska Center for Cross Cultural Studies.

Brown, H. Douglas. 2006. *Principles of Language Learning and Teaching*. Englewood Cliffs NJ: Prentice-Hall.

Demmert Jr., William G. & Towner, John C. 2003. *A Review of the Research Literature on the Influences of Culturally Based Education on the Academic Performances of Native American Students*. Portland OR: Northwest Regional Educational Laboratory.

Gay, Geneva. 2000. *Culturally Responsive Teaching: Theory, Research and Practice*. New York NY: Teachers College Press.

Goodkind, Jessica, Ross-Toledo, Kimberly, John, Susie, Hall, Janie Lee, Ross, Lucille, Freeland, Lance, Coletta, Ernest & Becenti-Fundark, Twila. 2011. Rebuilding trust: A community, multiagency, state and university partnership to improve behavioral health care for American Indian youth, their families and communities. *Journal of Community Psychology* 39: 452–477.

Grinevald, Collette. 2003. Speakers and documentation of endangered languages. In *Language Documentation and Description, Vol. 1*, Peter K. Austin (ed.), 35–51. London: SOAS.

Gruenewald, David. 2003. Foundations of place: A multidisciplinary framework for place-conscious education. *American Educational Research Journal* 40: 619–54.

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

Hinton, Leanne. 2003. "How to teach when the teacher isn't fluent. In *Nurturing Native Languages*, Jon Allen Reyner, Octaviana V. Trujillo, Roberto Luis Carrasco & Louise Lockard (eds), 79–92. Flagstaff AZ: Northern Arizona University.

Hinton, Leanne. 2013. *Bringing Our Languages Home*. Berkeley CA: Heyday.

Hinton, Leanne & Hale, Kenneth. 2001. *The Green Book of Language Revitalization in Practice*. San Diego CA: Academic Press.

Hymes, Dell. 2007. Languages and their uses. In *The First Oregonians*, Laura Berg (ed.), 245–262. Portland OR: Oregon Council for the Humanities.

Jacob, Roger. 2009. *Naktkwanint Tiinwitki (To Take Care of Things the Indian Way): Sustainable Management of Yakama Nation Forest Cultural Resources. A Final Report.* The Yakama Nation.

Jansen, Joana & Beavert, Virginia. 2010. Combining the goals of language documentation and language teaching. In *Building Communities Making Connections*, Susana Rivera-Mills & Juan Antonio Castillo (eds), 62–80. Newcastle upon Tyne: Cambridge Scholars.

Kinkade, M. Dale, Elmendorf, William W., Rigsby, Bruce & Aoki, Haruo. 1998. Prehistory: Introduction. In *Handbook of North American Indians,* Vol. 12: *Plateau*, William C. Sturtevant & Deward E. Walker (eds), 49–72. Washington DC: Smithsonian Institution.

Kolb, David & Fry, Ron. 1975. Toward an applied theory of experiential learning. In *Theories of Group Process*, Cary L. Cooper (ed.), 33–58. London: John Wiley.

Lipka, Jerry. 2002. *Schooling for Self-determination: Research on the Effects of Including Native Language and Culture in the Schools*. Charleston WV: Clearinghouse on Rural Education and Small Schools, AEL.

Mmari, Kristin N., Blum, Robert W. & Teufel-Shone, Nicolette. What increases risk and protection for delinquent behaviors among American Indian Youth? *Youth and Society* 41: 382–413.

National Education Association (NEA). 2010. *Focus on American Indian and Alaska Natives.* Washington DC: National Education Association of the United States.

National Education Association (NEA). 2011. *Voices of Native Educators: Strategies that support success of high school students.* Washington DC: National Education Association of the United States.

Nee-Benham, Maenette Kape'ahiokalani Padeken Ah & Cooper, Joanne E. 2000. *Indigenous Educational Models for Contemporary Practice: In our Mother's Voice. Sociocultural, Political, and Historical Studies in Education.* Mahwah NJ: Lawrence Erlbaum Associates.

Northwest Indian Language Institute (NILI). 2011. Northwest Indian Language Benchmarks. Ms.

Oregon Trout. 2005. *Salmon Watch Curriculum Guide*. Portland OR: Oregon Trout.

Rice, Keren. 2006. Ethical issues in linguistic fieldwork: An overview. *Journal of Academic Ethics* 4: 123–155.

Smith, Gregory. 2002. Place-based education: Learning to be where we are. *Phi Delta Kappan* 83: 584–594.

Sobel, David. 1996. *Beyond Ecophobia: Reclaiming the Heart of Nature Education*. Great Barrington MA: The Orion Society.

Supahan, Sarah & Supahan, Terry. 2001. Teaching well, learning quickly: Communication-based language instruction. In *The Green Book of Language Revitalization in Practice*, Leanne Hinton & Ken Hale (eds), 195–198. San Diego CA: Academic Press.

Thompson, Lawrence C. & Kinkade, M. Dale. 1990. Languages. In *Handbook of North American Indians,* Vol. 7: *Northwest Coast*, Wayne Suttles (ed), 30–51. Washington DC: Smithsonian Institution.

Viles, Jerome. 2010. Canoe journey as a model for language revitalization: Experiences with the Grand Ronde Canoe family. *University of Oregon McNair Scholars Program Research Journal* 1: 39–55.

Wiley, Judy. 2008. American Indian youth transition in Oregon: A strategic planning concept paper. Grants Pass OR: Oregon Department of Human Services.

Yamada, Racquel-María. 2010. *Speech Community-based Documentation, Description, and Revitalization: Kari'nja in Konomerume.* Ph.D. dissertation, University of Oregon.

_____. 2011. "Integrating Documentation and formal teaching of Kari'nja: Documentary materials as pedagogical materials". *Language Documentation & Conservation* 5: 1–30.

Zahir, Zalmai. 2007a. *A Plan for Teaching Lushootseed Language and Culture on the Canoe Journey*, unpublished ms.

_____. 2007b. *A Plan for Teaching Lushootseed Language and Culture in Schools,* unpublished ms.

_____. 2010. *Canoe Journey to Neah Bay Washington, 2010: A Lushootseed Self-immersion Experience*, unpublished ms.

Remembering ancestral voices

Emergent vitalities and the future of Indigenous languages

Bernard Perley
University of Wisconsin-Milwaukee

Language endangerment is a global tragedy that has prompted a surge in research and advocacy on behalf of those communities whose languages have been diagnosed as endangered. Indigenous languages in the Americas and Australia are the most at risk of becoming extinct by the end of this century. Graded scales from "safe" to "extinct" present diagnostic frames of reference that influence the kinds of approaches toward documentation and revitalization that community activists/advocates and language experts develop and initiate. Those languages deemed "extinct" and/or "severely endangered" are hampered by the prevailing metaphors that unduly constrain possible actions for language vitality. This paper offers a re-conceptualization of the metaphors regarding language endangerment away from "death" and "extinct" to "sleeping" and from documentation toward "emergent vitalities". This is especially critical for indigenous communities living in their ancestral homelands where *remembering* ancestral voices plays a significant role in possible futures for indigenous languages.

1. Introduction

This paper is based on the proposition- "If we are to save indigenous languages we must remember ancestral voices." This is not a simple proposition. Complexities are prevalent in determining who "we" represents, what interventions "save" might initiate, how we define "indigenous" and "languages", what forms "remembering" takes, and finally, who and/or what are "ancestral voices". I cannot promise to provide clear definitions, best practice interventions, or solutions for all endangered language scenarios. I can promise to offer perspectives that are both personally as well as professionally informed regarding endangered languages and language revitalization. My perspective draws from my personal experience of Native American language loss, my academic training in art, architecture, linguistic anthropology, as well as my commitment to enhancing the prospects for indigenous language vitality. Rather than view language documentation as desperate measures to

record the last words of the last speakers, I re-conceptualize language documentation as alternative vitalities that can potentiate indigenous emergent vitalities. To that end I offer a theoretical and practical stance toward language documentation and revitalization that *remembers* indigenous ancestral voices as timeless resources for all the meaningful relations that integrate stories, landscapes, spirituality, and relationships as living indigenous worlds.

1.1 *Remembering*, embodiment, vitality, and indigenous voices

I am using the terms "remember" and "embody" and their derivatives to provoke alternative interpretations to put forward my argument on "*remembering* indigenous voices". I italicize *remember* and its derivatives to evoke the cognitive act of bringing forth from memory as in remembering vocabulary, concepts, etc. as well as to highlight the act of reincorporating an entity that once enjoyed membership within a group back into that group. This is significant because this action reverses the process of "dismemberment" that linguistic science unwittingly inflicts on indigenous languages. I argue that the linguistic practice of reducing languages down to their constituent parts such as phonemes, morphemes, syntax, and semantics may on one hand provide insights for linguistic science but on the other hand it also *dismembers* the subject language. This practice *dismembers* language in two ways; first, it breaks the language apart into formal analytical categories and second, it *dismembers* the language from the community of speakers. The first part has the unintended consequence of limiting speakers of subject languages to the role of informant. Their contribution to linguistic science is that of data source. The professional elicitation methods, documentation practices, and subsequent analyses are focused on the linguistic code and speech events, not the collaborative relationship between native-speaker-consultant and scholar. This leads to the second aspect of *dismemberment* that is especially significant for indigenous languages. *Dismemberment*, interpreted as discontinued membership within a community, is one of the most devastating factors in the endangerment of indigenous languages. For example, in native North America the phrase "all my relations" speaks to the relationships native peoples have to their extended community, including their heritage landscapes. Formal linguistic analyses and documentation of indigenous languages become *dismembered* from indigenous communities in both senses of the term, separated into constituent parts and separated from the community of speakers. These forms of *dismemberment* also contribute to the disembodiment of indigenous languages.

The "death" of a language is generally considered the point at which "nobody speaks it anymore" (Crystal 2000: 1). Crystal also states, "to say that a language is dead is like saying that a person is dead. It could be no other way – for languages

have no existence without people" (ibid.). I agree, but only to a certain point. I have argued that language loss can be viewed as increased disembodiment (Perley 2011). The many factors for reduced language use by speakers within a community correspond to the growing silence of the indigenous language within the community. When the indigenous language is no longer used it can be inferred to have "died". For language to have "life" it must be spoken by speakers. Language life, then, is language embodiment in the speaking community. However, I am uncomfortable with the language life and death metaphors. I prefer to use the term "vitality" to describe language ontology.

The biological metaphor of language life is a powerful and relatable imagining of the relative state of a language. UNESCO's interactive online *Atlas of the World's Languages in Danger* (Moseley 2010) emphasizes nine criteria to evaluate the relative "vitality" of any given language. Once UNESCO's consultants declare a language as "endangered" they designate the following relative degrees of endangerment- vulnerable, definitely endangered, severely endangered, critically endangered, and extinct (accessed 1-30-2013). Notably, the UNESCO calculus of vitality presents those languages that enjoy greatest vitality as "safe" which is defined as "language is spoken by all generations; intergenerational transmission is uninterrupted" (accessed 1-30-2013). After "safe" there are only degrees of endangerment culminating in extinction. I have argued that the biological metaphor has its limitations to point out the unintended consequences that such metaphors present for language advocate/activists whose indigenous languages have been designated as extinct as well as those languages perceived as endangered (Perley 2011, 2012a). Assessments of vitality on a graded scale are important diagnostic rubrics but I prefer to emphasize vitality as processual rather than diagnostic. I use "vitality" to evoke the activity and energy of emergent states in addition to life sustaining energy. This shift in ontological conceptualization permits the inclusion of non-living items such as lexicons, grammars, DVDs, etc, to be afforded "energy" for sustaining language life. These material artifacts of language analysis and documentation are important life giving sources for language revitalization and awakening. These artifacts are only language life affirming when used by language speakers and/or learners. I refer to this contingent ontology as "alternative vitality" to expand our perception of what constitutes language life. Alternate vitalities as a conceptual stance imbue material artifacts with language life energy. More critically, the interaction between these material artifacts of language and speakers/learners increases the vitality of language through their intersubjective relations. This is crucial for indigenous language revitalization and especially true for "dead" or "extinct" languages. The expanded field of intersubjective relations, including the heritage landscape, makes it possible to reverse linguistic *dismemberment* and disembodiment of indigenous languages. This active process of *remembering* and

re-embodiment is necessary for proposing "emergent vitalities" as a perceptual stance for revitalizing indigenous voices.

The term "indigenous" is as complex as it is imprecise. There are growing numbers of publications that purport to explain indigenous experiences (Maaka & Anderson 2006; de la Cardena & Starn 2007), indigenous methodologies (Denzin et al. 2008; Minde 2008; Smith 2005 [1999]), and the politics of indigeneity (Niezen 2004; Stewart-Harawira 2005; Maaka & Fleras 2005; Shaw 2008). In this paper I use the term as a general reference to autochthonous populations who articulate original inhabitation of their heritage landscapes. The term 'aboriginal' will be used interchangeably with First Nations when referring to the indigenous communities in Canada. The terms Native American and American Indian are equally complex and imprecise and I will use one or the other as the context requires. This paper distinguishes between indigenous languages and minority languages for a couple of reasons. First, my approach for indigenous language revitalization requires *remembering* ancestral landscapes. Second, many indigenous languages may be categorized as minority or minoritized languages but that is not always the case. Nor are all minority/minoritized languages categorized as indigenous. Also, there are minority languages that are not considered endangered. I am aware that categories such as indigenous and minority are not mutually inclusive or exclusive, static or immutable, and are subject to change in response to contingencies of emergence (Perley 2009). This paper accepts all the complexities of such categories because my perceptual stance for indigenous language revitalization encourages those emergent vitalities of indigenous voices.

1.2 Language crisis and intervention management

There is no question language endangerment and subsequent language extinction is an immediate and tragic crisis. The turn of the millennium has been a pivotal period of linguistic activity that has prompted dramatic increases in publications, program building, and intervention planning dedicated to recognizing the tragedy of language death, diagnosing language endangerment, implementing language documentation, and developing best practices for language conservation and/or revitalization. There is a corresponding increase in the proliferation of conferences, symposia, and websites that convey the same critical message- the world's languages are dying at an alarming rate and if nothing is done to document them we will lose a significant portion of our human heritage. The call to action is urgent and justifiable. Projected rates of language loss vary from "over half of the languages of the world" (Hinton 2001a: 3 quoting Krauss 1992) to as much as ninety percent (Grenoble 2008: 216, in Austin 2008) of the world's languages becoming extinct by the end of this century. The National Geographic Society has produced publications such a K. David Harrison's *The Last Speakers* (2010)

and a recent magazine article (Rymer 2012). The Society has also developed their website, *Enduring Voices,* dedicated to calling attention to the crisis. As you access the website you are greeted with a blue screen in which a quivering white line forms the sentence "Every 14 days a language dies".[1] The blue field then turns into a global map identifying the language "hotspots" defined as geographic locations where "many languages near extinction" are clustered (National Geographic Society, accessed 9/17/2012). Such graphic, textual, and interactive media attempt to convey to general reading, watching, and listening audiences the gravity of the situation as well as the need to support efforts to forestall the extinction of the world's languages.

Significantly, this is not only a linguistic crisis. The communities whose languages are imperiled will be directly affected by continued language attrition and eventually language death. The extinction of language is also the extinction of culture and identity as well. These collateral extinctions (Perley 2012b) are critical reminders that any language documentation and revitalization effort must include the community of speakers as well as their cultures and identities. In short, language survival necessitates cultural and identity survival.

Personally and professionally I acknowledge and appreciate the important contributions linguistic science and related fields have made to indigenous language analysis, documentation, and resultant conservation and/or revitalization programs. There is a great deal of inspiring work done in the growing field of linguistic documentation. Nikolaus Himmelmann defines the aims of documentary linguistics as providing "a comprehensive record of the linguistic practices characteristic of a given speech community. Linguistic practices and traditions are manifest in two ways: (1) the observable *linguistic behavior*, manifest in everyday interaction between members of the speech community, and (2) the native speakers' *metalinguistic knowledge*, manifest in their ability to provide interpretations and systematizations for linguistic units and events" (Himmelmann 1998: 166). This approach emphasizes the difference between documentation for analysis as part of a descriptive linguistic practice and documentation as a record of linguistic practices of a given community. Himmelmann characterizes this distinction as the primary focus such that the primary data can serve any research interest. Equally significant is Himmelmann's insistence that documentary work be participatory with the rights and interests of the subject community as an ethic and moral

1. See ⟨http://travel.nationalgeographic.com/travel/enduring-voices/⟩. The website is part of the collaborative effort between the National Geographic Society and the Living Tongues Institute to promote awareness of the language crisis but also to promote the maintenance, preservation, and revitalization of endangered languages ⟨http://www.livingtongues.org/⟩ (20 October 2012).

obligation. Along with Himmelmann's formulation of the field of documentary linguistics there have been parallel conversations covering theoretical concerns (Dorian 1989; Grenoble & Whaley 1998; Mühlhäusler 1996, 2003; Himmelman 1998) to technological innovations (Bird & Simmons 2003; Eisenlohr 2004; Penfield, Cash, et al. 2006; Berez & Holton 2006) to best practice interventions (Hinton & Hale 2001; Grenoble & Whaley 2006) to multimodal documentation and analysis (Himmelman 1998; Jewitt 2011 [2009]; Csató & Nathan 2003) just to name a few. All these examples underscore the serious nature of language endangerment as a global crisis. They also reflect the complexity of diverse cases requiring innovative strategies for intervention and documentation based on contextually significant parameters. Echoed across these important contributions is an influential rhetoric conceptualizing language as a biological organism. Language "death" and "extinction" are tragic endpoints for endangered languages. Documentation and conservation efforts were initiated to "save" endangered languages from death and extinction. The situation is not so easily ameliorated. Such interventions, beneficial as they are, require a critical evaluation of the "endangered biological organism" metaphor used by language experts and popular media to describe language endangerment. Doing so will reveal how influential this conceptual framing is for promoting particular kinds of expert interventions.

The rhetoric of language endangerment is a powerful discourse strategy that language experts use to inform non-specialists of the gravity of the language crisis in terms that are relatable but also problematic (Hinton 2001b; Nettle & Romaine 2000; Harrison 2010; Perley 2011, 2012a). Furthermore, language life versus language death invokes ontological language states that unnecessarily constrain language advocacy in many indigenous language communities whose language vitality is described variously as endangered, severely endangered, moribund, or extinct while ignoring potential solutions to indigenous language endangerment (Perley 2011, 2012a). I argue that as language experts and advocates we must shift our metaphorical conceptualization away from the death and extinction metaphors toward a conceptualization that provides the conditions for language life. I suggest "emergent vitalities" as an operational metaphor for our expert analysis and subsequent intervention. Shifting to such a conceptualization places the community of speakers, their cultures, and their identities as mutually interdependent aspects of indigenous languages that emphasizes possible futures for indigenous language, culture, and identity.[2]

---

2. My stance is informed by antecedent scholars such as Franz Boas 1966 [1940], Edward Sapir (Mandelbaum 1985 [1949]), and Benjamin Lee Whorf (Carroll 1970 [1956]) who made similar arguments in the early twentieth century. It is echoed in more recent research programs such as cultural ecology (Steward 1972; Bateson 2000 [1972]), humanistic geography

2. Metaphors of language ontologies

Today, the Maliseet communities are faced with the specter of language death (Perley 2011). Maliseet is an Eastern Algonquian language that is spoken in Eastern Maine (USA) and along the St. John River in New Brunswick (Canada) and it is often hyphenated with Passamaquoddy as the two languages are mutually intelligible (Erickson 1978; LeSourd 2007). Maliseet-Passamaquoddy language vitality has been described as "severely endangered" on the UNESCO online atlas of *World's Languages in Danger* (Moseley 2010). The online atlas designates a language as "severely endangered" when the "language is spoken by grandparents and older generations; while the parent generation may understand it, they do not speak it to children or among themselves" (Moseley 2010).[3] My own fieldwork and the resultant ethnography (Perley 2011) confirm the UNESCO assessment is an accurate description of the state of the language for Tobique First Nation.[4] Furthermore, in my ethnography I describe the various factors which, when taken together, constitute processes that continue to undermine the Maliseet language through *dismemberment* and disembodiment from the community of speakers. Factors such as social upheaval, which include resettlement, decreased isolation, urbanization, and increased intermarriage; assimilatory pressure such as media, education, stigma, cultural hegemony; and changes in values reflected in the lack of literature, reduced or loss of language use, reduction in number of speakers and most important of all, changes in attitude of the community toward their indigenous language (Perley 2011: 43–44).[5] Not all factors were present for the Maliseet case at Tobique First Nation and not any one factor was the determining factor for the Maliseet language shift toward "severe endangerment". I described how those factors contributed to the gradual *dismemberment* and disembodiment of the Maliseet language from the Maliseet community. Each factor represents the loss of a particular domain of language use in the community; or, *dismemberment*.

(Tuan 1990 [1974], 1977, 1979), linguistic ecology/eco-linguistics (Mühlhäusler 1996, 2003), linguistic documentation (Harrison 2007; Evans 2010), linguistic anthropology (Basso 1996; Thornton 2008; Perley 2011), and many more related fields and interdisciplinary programs. The limitations of space prevent a more complete review of relevant programs and scholars.

3. It should be noted that you may also find Maliseet spelled *Malecite* and hyphenated with Passamaquoddy to read *Malecite-Passamaquoddy;* see Ethnologue at the following –⟨http://www.ethnologue.com/language_index.asp?letter=M⟩ (21 September 2012).

4. I have been given permission by the Chief of Tobique First Nation to use the name of the community in my publications.

5. These "factors" were drawn from *The Loss of Australia's Aboriginal Language Heritage,* Annette Schmidt et al. (1990). Additional diagnostic rubrics can be found in Edwards (1992), Grenoble and Whaley (1998), Nettle and Romaine (2000), and Tsunoda (2006).

Consequently, from the perspective of experts in language endangerment the Tobique First Nation community is experiencing the linguistic ontological shift of their indigenous language from "living" to "dead"; or, disembodiment.

Even though the biological metaphor is understood to be problematic it is used to assert that "languages are intimately connected with humans, our cultures, and our environment" (Nettle & Romaine 2000: 6) as well as finding a way to accurately "describe people abandoning complex systems of knowledge like languages" (Harrison 2007). These authors, Nettle and Romaine as well as Harrison, articulate the complex and intimately connected knowledge systems that languages represent and the concomitant collateral endangerment entailed. Unfortunately, language scholars are often influenced by the metaphor to conceive of language endangerment and language death as death of the linguistic code rather than the death of the social relations language mediates. This conceptual framing results in two shortcomings. First, the living/dead dichotomy precludes other ontological states for language. Second, language is conceptualized as an object that can be taken apart, analyzed, and documented. These two shortcomings propagate unintended constraints on the critical work of language revitalization in indigenous communities. Expert interventions operating from this conceptual stance have an objective of saving the Maliseet language through documentation practices. The unintended consequences of focusing on the production of artifacts of living language will render Maliseet a language that is *dismembered* and disembodied from the community of speakers. For example, a dictionary entry for the Maliseet culture hero is as follows-

(1) **Koluskap**. *noun animate*. Name of first man, according to Wabanaki oral tradition (in English, often spelled Glooscap, Gluskap, etc.).
(Francis & Leavitt 2008: 196)

We learn the culture hero's name, it is an animate noun, and the name comes from Wabanki oral tradition but we learn nothing of the context or the significance of either the name or the culture hero. Compounding this dismemberment is the omission of the cultural polysemy *Koluskap* represents. *Koluskap* has been translated as "liar" (Ives 1964: 17) and is often used to denote a person who lies as well as characterizing the culture hero as a liar. In the same dictionary mentioned above, the entry for "liar" is-

(2) **koluskapiyiw, koluskapihiw**. *verb ai 3(2).* s/he is a liar, tells tall tales, exaggerates, bluffs; s/he is deceitful. (Francis & Leavitt 2008: 196)

Again, there is minimal information provided and it is not contextualized in the everyday polysemic imaginations and usages across generations of Maliseet community members. The English word *Liar* has serious implications but *Koluskap*

can simultaneously invoke serious as well as humorous Maliseet social relations (Perley 2012c: 184). The dictionary entries do not allow us to understand the social/cultural context, or the experiential polysemy of *Koluskap*. The dictionary further disembodies Koluskap through textual capture, relegated as a reference, and "saved" for later access.

Despite such *dismemberment*, documentary linguistics represents potential for meaningful revitalization. The growing field of documentary linguistics is crucial for recording the wealth of linguistic data from endangered languages in the hopes that such data will provide significant insights for linguistic science but, as Himmelmann insists (1998), the documentation products should also serve to benefit the communities from which those documents were produced. Documentation may preserve in concrete form evidence of spoken ephemera of conversations and other speech practices but it is not the same as providing the conditions for the maintenance and renewal of speech as a mediator of social relations. In order to realign documentation efforts toward enhancing the social relationships, a realignment of the metaphors we language experts use is necessary. The concepts of 'sleeping languages' and 'emergent vitalities' are two perceptual stances language advocates/activists take toward rethinking ontological states of language to better identify potential actions to "awaken" indigenous languages and/or potentiate indigenous language vitality.

2.1 Emerging from sleep

Extinction is generally regarded as a permanent state of non-existence. But, is that true for languages? Can an extinct language be revitalized? There are notable examples of language experts, advocates, and activists who offer hopeful answers to these questions. These community language advocates/activists have critically evaluated the discourses of language death and extinction and rejected the metaphor in order to conceive of alternative vitalities for their indigenous languages. I use the term "alternative vitalities" to ascribe subjective qualities to the artifacts of spoken language (such as lexicons, grammars, storybooks, CDs, DVDs, audio and video recordings, etc.) that document endangered as well as "extinct" languages (Perley 2011). I do so for a couple of reasons. First, documentation practices too often privilege the linguistic code while relegating the speakers to the status of conveyors of linguistic data to serve linguistic science. In these cases, documentation does not constitute living language. They are artifacts of live exchanges between speakers and therefore the subject language is not "saved" (Perley 2012a). Second, if we view all forms of language – documents and discourses–as subjectivities we can set aside the death and extinction descriptors and posit the possibilities of language life for those communities whose languages are diagnosed as dead or extinct.

Leanne Hinton, commenting on the prevalence in the literature to describe languages without speakers as "dead" or "extinct", states "I prefer the less final metaphor of "silence," or L. Frank Manriquez's "sleep" (Hinton 2001b: 413). Significantly, in another chapter discussing the various states of language endangerment, Hinton quotes Frank Manriquez, "(d) Languages that have lost all their speakers, so that the only record of them (if any) lies in notes and recordings by linguists. These are beyond being "endangered" and are usually called "dead languages," although L. Frank Manriquez, whose ancestral language, Togva, has no speakers, "prefer[s] to think of them as sleeping" (personal communication)" (Hinton 2001a: 4). The critical difference between "dead" languages and "sleeping" languages lies in the possible interventions language advocates/activists can initiate. Specifically, language death is a final ontological state precluding any possibility for intervention whereas sleeping languages presumes language life and therefore potentiates the prospects for awakening the language.

Daryl Baldwin of the Miami Nation of Oklahoma has taken the "sleeping" metaphor and used it to "awaken" the Miami language. Baldwin is featured in the Miami Nation of Oklahoma video *myaamiaki eemamwiciki, Miami awakening* (2008) wherein he states that from the biological perspective "extinct means gone forever." What does that mean for the Miami language, a language that linguists had declared "dead" and/or "extinct"? Baldwin states, "maybe these academics are wrong. Maybe we can reconnect. Maybe Miami can be the language of emotion, the language of thought, not just the language of speech" (2008). These thoughts led Baldwin to initiate a Miami language awakening project in which he coordinated a collaborative project with Miami University of Ohio. Baldwin, together with the Miami Nation of Oklahoma and Miami University established the Myaamia Project. Those efforts are now extremely important in the awakening of the Miami language. The shift in metaphor from "extinct" to "sleeping" was key to making Miami the language of emotion, thought, as well as speech.

In the Miami case, it was not enough to change the metaphor. The distinct advantage that Baldwin had in his efforts to awaken his ancestral language was access to extensive documentation of the Miami language. The historical documents provided the materials needed for lexicons, grammars, and social/cultural knowledge. Baldwin, in collaboration with linguist David Costa, is awakening the Miami language (Baldwin & Olds 2007). Those documents are historical artifacts of colonial, church, and government relations but they do provide records of the conversations that were taking place at a particular place and a particular time in Miami social/linguistic history. Even though these artifacts represent language as object and therefore available for analysis, Baldwin saw the documents as the catalyst for awakening the Miami language. Rather than view the documents as inert lifeless objects, Baldwin was able to breathe life into the Miami language through

those documents. In this way there was a subject to object relationship that suggests alternative ways of "being" for Miami language documents. The critical point here, as I argued in 2011, is- "the texts and recordings may be inert, but they also hold great potential for revitalization, or, re-embodiment. The inert language entities can become intersubjective catalysts for individual subjects, or for a number of subjects, if the context of their animation will allow it" (Perley 2011: 147).[6] For sleeping languages, as in the Myaamia case, the documents can prompt individuals or communities to read, listen to, or speak through them. Those activities, as well as acts of producing texts, audio and video recordings, and conducting language classes, are all active linguistic processes indicative of language vitality. Historical documentation of sleeping languages, then, should not be viewed as having "saved" those languages for later analysis or as a record of linguistic behavior of a given language. Rather, documentation, as Baldwin illustrates, is an opportunity to imbue those documents with alternative vitalities thereby allowing community members to creatively participate in the emergent vitalities of everyday communication. Sleeping languages are not the only languages that can benefit from re-conceptualizing documentary artifacts as alternative vitalities.

2.2 Emergent vitalities

There is no doubt many cases of extreme language endangerment require documentation strategies to preserve linguistic knowledge of endangered languages. The products of documentation practices are directed toward recording and providing linguistic data for linguistic science but those documents may also provide opportunities for awakening languages for a later generation as illustrated in the Myaamia case. In those cases where community members are awakening their sleeping language the resource of extensive documentation is critical for language revitalization efforts. Can the same be said for languages considered endangered or severely endangered?

The aim of documentary linguistics, as outlined by Himmelmann, is "the record of THE LINGUISTIC PRACTICES AND TRADITIONS OF A SPEECH COMMUNITY" (emphasis in the original, Himmelmann 1998: 166). This entails documenting communicative events as "part of the larger communicative setting"

---

6. The intersubjective relations quoted in this excerpt refer to the phenomenological argument presented in chapter one of my ethnography. Limited space prevents me from going into detail in explaining this argument but briefly stated, the experiential aspect of language is grounded in how the subject (or individual) must engage in conversation and/or any other communicative activity with other interlocutors, or subjects. The communicative relationship is intersubjective, even when one "talks" to ones self.

(ibid:168) and providing metalinguistic commentary from community members. For indigenous language advocates/activists, among the most promising aspects of Himmelmann's program for documentary linguistics is his insistence that the researcher consider the wishes of the community.

> "How can a documentation project be presented to a community in such a way that the community is likely not only to accept it but also to shape it in essential aspects? In some communities there may exist some strong ideas about how documentation should proceed, which do not accord with the researchers' plans. How can and should these conflicts be resolved? Closely linked to the issue of participatory design of a documentation project is the issue of the researchers' involvement in language-maintenance work which may be of greater interest to the community than just a documentation." (Himmelmann 1998: 188–189)

These considerations are laudable and promising but I point out that Himmelmann's project is to "conceive of language documentation as a field of inquiry and research in its own right" and "that language documentation is NOT some kind of "theory-free" enterprise. Instead, documentary linguistics is informed by a broad variety of theoretical frameworks and requires a theoretical discourse concerned with conceptual and procedural issues in language documentation" (ibid:190). I recognize the value of the program as a field of linguistic research and inquiry that produces artifacts of communicative events through holistic documentary practices. I am concerned that the conceptual emphasis is focused on the "record" as the end product. Yes, the artifacts can be used for other purposes, be it linguistic analysis or language maintenance, but I would like to re-conceptualize the emphasis away from documenting communicative events to cultivating social relationships where documentation happens as emergent vitalities. Placing emphasis on the social relations permits the inclusion of all social relations in many indigenous communities who are working to *remember* their ancestral landscapes, stories, beliefs, and experiences. As language scholars working on language documentation projects, perhaps we can initiate relationships with indigenous language advocates/activists so that we *remember* all those social relations. Doing so will introduce us to the voices of the ancestors.

3. *Remembering* ancestral voices

The native language class in Mah-Sos school (on Tobique First Nation) had a regular storytelling exercise. The native language teacher would tell the students that they were going to hear a traditional Maliseet story. The students were always excited about stories because they were allowed to draw their own illustrations

for the stories. As all the students settled down to listen to the story the native language teacher would open a large format book and turn to the appropriate page. She would read a story about the exploits of the Maliseet culture hero Kluskap as he shaped the world into its current form.[7] The stories were short, they were read in English, and they were at various times in the historical documents considered Algonquian, Abenaki, Wabanaki, as well as Maliseet stories (Perley 2012c). These aspects of the storytelling exercise have serious implications for the devaluation of oral traditions and Maliseet language extinction (Perley 2012c) but for the purposes of this paper I focus on the prospects of linguistic, cultural, and spiritual repatriation of Maliseet stories, language, and landscapes. In short, Maliseet worlds.

3.1 New domains, new relationships

In the community of Tobique First Nation there are many independent projects that go unacknowledged but represent the creative linguistic vitality of the community in producing new domains of Maliseet language usage. Translations of popular songs such as *Rudolf the Red-nosed Reindeer* or the Beatle's *Yellow Submarine* as well as *O Canada* and *Silent Night*, translations of popular cartoons such as *Peanuts* and the *Wizard of Id*, word search puzzles, and many digital video projects of community members engaged in traditional practices such as making baskets, braiding sweetgrass, and gathering fiddle-heads are some of the many projects that I am familiar with. I am certain there are many more projects and activities that escape my attention. Several collaborative projects have been initiated by members of the community that produced translations of children's books and subsequent recordings of the Maliseet speakers reading the text and recording them onto CDs. These documentary practices are records of Maliseet speakers who draw from products of the dominant popular culture and recast them into the Maliseet language as a creative expansion of Maliseet linguistic domains. These examples display creative adaptation of non-Maliseet cultural products but it should be noted that not all practices are drawn from outside the Maliseet

7. The storytelling event was typical of many such storytelling activities I observed when doing field research in Mah-Sos school on Tobique First Nation during the mid nineteen-nineties. Please see Perley 2011 for a discussion of the unforeseen conundrums in the classroom as the native language teacher shifted from oral-based pedagogy to text-based pedagogy. Briefly, the language teacher assumed that she learned the language through oral transmission as a natural way to learn language. Also, she wanted to avoid using any text to teach the children the language because there was no orthographic standard for representing Maliseet. However, when the students failed to retain orally taught Maliseet words, expressions, etc, the teacher decided to use written forms of the same exercises to help the students retain Maliseet.

culture. For example, (a) the tribal government has collaborated with documentary film producers to produce a DVD featuring Maliseet speakers engaged in a variety of activities where the language is actively used, (b) adult Maliseet language classes are offered intermittently, and (c) Maliseet speakers collaborate on producing online dictionaries. These examples are drawn from local Maliseet worlds but they have the potential to reach audiences beyond the community of Tobique First Nation.

One significant project that has propelled the Maliseet language from being an Eastern Canadian aboriginal language into a nationwide language is the "versioning" project produced by community member Jeff Bear (executive producer for Urban Rez Productions).[8] Jeff Bear and Marianne Jones (producer/director) have produced three television series *Ravens and Eagles*, *Storytellers in Motion*, and *SAMAQAN: Water Stories*.[9] All three productions are intended for the television market and are significant for their content and the mission driving their production. Their mission, as expressed on their website, is – "UrbanRez.ca is committed to storytelling in the digital media industry. Our primary goal is to find relevant, timely and inspiring stories of the life and times of Canada's aboriginal community."[10] The importance of that mission cannot be underestimated. The ability to connect disparate aboriginal communities with inspiring stories that in turn inspire other creative aboriginal people to produce timely and relevant content in all forms of media is an example of "alternative vitality" in practice. Even more significant is the parallel project of creating Maliseet language voice-overs for the *Storytellers in Motion* and *SAMAQAN: Water Stories* series. The project requires taking the scripts from the original shows to produce a parallel script that expresses the same ideas in Maliseet. It would be an error to describe the process as translation. The producers enlist a number of Maliseet speakers to provide a Maliseet voice for the people in the documentaries. During the process new words and expressions had to be negotiated between the producers and the speakers. The end result was the production of versions of the original series that have Maliseet language voice–overs accompanied by English captions. Jeff Bear states that among the most significant aspects of these projects are the great number of recordings of Maliseet language use in the modern context, the

8. I wish to thank Jeff Bear for reviewing and approving this segment of the paper and giving me permission to print this segment as well as allowing me to use his name and the name of his production company (September 21, 2012).

9. Please visit the website- ⟨http://www.urbanrez.ca/-⟩ for more information.

10. Accessed September 21, 2012.

development of new vocabulary terms that capture the "contemporary Maliseet and aboriginal world",[11] and the broadcasting of the shows across Canada.

These projects were possible because Jeff Bear was raised speaking the Maliseet language. This gave him the skills to narrate some of the episodes as well as "write" the Maliseet scripts in Maliseet. As Jeff Bear explains, "It's not like I sit and write words. The orthography is diverse, not standardized and complex. It is easier to take an oratorical approach. I sit and contemplate what to say and go into the booth and speak the story. Sometimes I'll check it with Shirley or Henrietta. And there are times when I feel I don't need to. It depends on each context."[12] His technical skills in writing and directing allowed him to tell the Maliseet version of the stories in this new medium. Furthermore, during the critical first year of production, Jeff Bear was the recording technician and made direct contributions to creating new terminology for modern contexts.

These television productions represent new media, new domains, and new linguistic practices for the Maliseet language. These are important contributions but I argue the most important factor that made the project possible is the bringing together of Maliseet speakers. When Jeff Bear gathered the Maliseet speakers together to participate in the project he also created a new context for Maliseet language use as well as creating new social relations among the speakers. Having witnessed one recording session, I was able to see the delight in the participants as they used Maliseet to converse with one another, negotiate proper terms for new vocabulary, and develop or strengthen friendships because of their Maliseet language use. It was a clear example of an emergent vitality for a living Maliseet world.

The above examples illustrate the creative language use by Maliseet community members. Those activities also reflect significant contributions to language vitality by establishing new language domains and new socio-linguistic relationships. I am encouraged by the creativity of the Maliseet community in finding new domains for language use but I want to turn to a key insight that Nancy Dorian shared when she described East Southerland Gaelic language endangerment. She argued that language tip toward obsolescence was directly associated with the loss of prestige for East Southerland Gaelic (1981: 52). When Gaelic elites switched to English it devalued Gaelic but more importantly it devalued the person speaking Gaelic. This loss of prestige was a key factor in undermining the status of Gaelic within the community. The same can be said of Maliseet in the largely English/French dominant culture of Canada. I argue, however, that language tip need not be in the

---

11. Personal communication 9/24/12.

12. Personal communication 9/25/12. Shirley is Jeff's sister and Henrietta is my mother.

direction toward obsolescence and resultant extinction. I argue that indigenous languages can tip toward vitality. The above examples display creative uses for the Maliseet language in new domains and bring prestige back to the Maliseet language. I argue that as a member of the Tobique First Nation community we can also find creative ways to *remember* the Maliseet stories, language, landscape, and spirituality. Doing so will instill more prestige back into the Maliseet world and tip the language toward vitality.

3.2 *Remembering* language, landscape, and spirituality

Remembering language, landscape, and spirituality is not just about the cognitive process of memory recall. I use the term to also indicate the process of *remembering* the social relations that linguistic science has *dismembered* from indigenous languages. After centuries of *dismemberment* by colonial processes, indigenous communities are poised to *remember* their languages with their landscapes and spirituality through their stories. This stance recognizes the kinds of subjectivities and social relations in indigenous communities that are not limited to interactions between humans. The phrase "all my relations", as used in native North America, is not restricted to siblings, parents, grandparents, cousins, community members, and strangers. The concept of social relations extends to the living world that indigenous languages allow speakers to intersubjectively co-inhabit. The following is a brief discussion of projects that I have developed to rethink language documentation practices and the *rememberment* of all Maliseet relations.

"Askəmi Pihce"[13] (A long time ago) is the first sentence in a graphic novel I am producing that links the Maliseet language with Maliseet oral traditions, the ancestral landscape of the Maliseet at Tobique First Nation, and spiritual aspects of Maliseet indigenous experience; in short, it is *remembering* Maliseet ancestral voices. *Apotamkon*, the graphic novel, can be broadly interpreted as a documentation project that preserves the Maliseet story about the Tobique Rock. My focus, however, is not on preserving what is left of the Maliseet language before it becomes extinct but one that invites readers to inhabit the living world of the Maliseet language and thereby renew the language. The graphic novel is also designed to reverse the long process of linguistic colonialism that has *dismembered* the Maliseet language from the community of speakers over the last four hundred

13. This spelling represents the orthographic practice of my Maliseet language teacher and consultant, Henrietta Black. Members of the community, as well as the scholarly community, recognize her formal knowledge of the language, her skillful use of the language in all domains of usage, and her translation/interpretation skills between Maliseet and English. She is also my mother.

years (Perley 2012c). Finally, the most important aspect of the project is to create the conditions for enhancing possible futures for the Maliseet language. It is an example of emergent vitality. For example, I described how the dictionary entry for *Koluskap* was an illustration of *dismemberment* of the Maliseet language from the Maliseet community. In the graphic novel I locate the story in our traditional landscape, tell the story so that the traditional story is meaningful for contemporary audiences by addressing contemporary concerns while re-integrating the many facets of Koluskap's character. Koluskap is not merely "the first man" or a "liar" but he is represented as storyteller, transformer, trickster, creator, and a wise man (see Figure 1). My version *remembers* those critical social relations that define Maliseet worlds as grounded in our oral tradition, as we experience it today, but also to enhance the prospects for Maliseet futures.

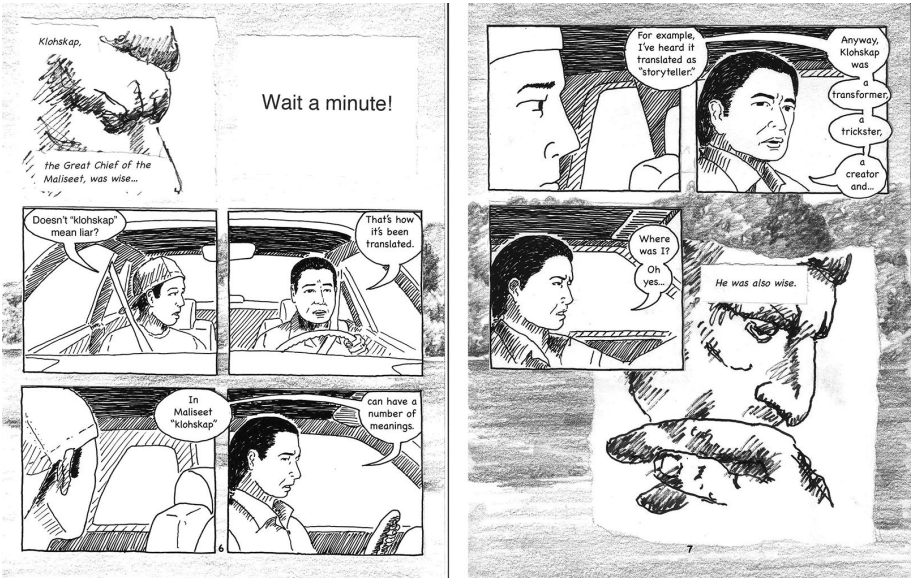

Figure 1. The many facets of Koluskap

As a graphic novel, *Apotamkon* is a contemporary telling of the Tobique Rock story. It is forty pages long; twenty pages are in Maliseet and 20 pages are in English. The illustrations utilize graphic novel format in the arrangement of panels on the page. Each panel narrates the story in textual as well as graphic form. Some of the innovations in graphic novels that are echoed in *Apotamkon* are the use of different fonts to suggest different voices, varying the size of the panels to indicate different narrative strategies, and developing a popular medium for broader reader interest. These graphic strategies of storytelling have become increasingly

popular as more "literary" texts such as the works of Jane Austin, Shakespeare, Ray Bradbury, Philip K. Dick, etc. are re-presented in graphic novel form. Despite the broad appeal of the graphic novel format, I recognize that there are some shortcomings in this medium as it is currently conceived. Its approach and its appeal is conceived and consumed as entertainment. My approach is to take advantage of the consumptive and entertainment aspect and turn it toward Maliseet language, cultural, and spiritual repatriation. How might that work?

Apotamkon invokes ancestral voices for present day concerns and anticipates lessons for the future. This approach is a deliberate break from the established practice of collecting Native American stories and publishing them as children's stories, just so stories (fanciful stories explaining natural phenomena), and/or folklore collections. Rather than a collection or retelling that is *dismembered* from the community of origin, my retelling is imbedded in the Maliseet landscape, the Maliseet community, and Maliseet spirituality. For example, the illustrations depict the landscape in the immediate environs of Tobique First Nation. Community members can identify the landmarks and the landscape in personal experiential terms. Moreover, the story is based on Maliseet oral traditional stories that describe how the landscape was formed from "time immemorial". The content of the story as textually represented in the dialogue and narration is augmented by the images that distinguish between temporal and experiential frames of reference. Small frames correspond to the temporal constraints of the fast pace of contemporary life while the larger frames cut across two pages to reflect the larger temporal frames of "deep time" (Perley 2002). The background of the frames is a reconstruction of the 360 degree panoramic landscape as viewed from the legendary Tobique Rock. This is significant because the novel is designed to be unfolded as you read it rather than turning pages from page one to page twenty. The reader can unfold the twenty pages of the story in Maliseet (English is on the obverse side) and link page one to page twenty and be surrounded by the Maliseet primordial landscape. Not only can the reader read the story but s/he can experience the story in the Maliseet language, in Maliseet space, and in Maliseet deep time. *Apotamkon*, while drawing from the graphic novel form, is less a novel and more a storytelling experience. Most importantly, *Apotamkon remembers* the Maliseet language with the traditional stories, the local landscape, and local spirituality through the embodiment of story, place, and experience. The graphic novel brings Maliseet oral traditions into a popular contemporary form of storytelling. Rather than conform to the expectations of the graphic novel genre, my adaptation breaks the frame of "reading" in order to re-embody the Maliseet language into Maliseet lives. The immersive storytelling experience is not just about "reading" the story but also about experiencing the story.

The *remembership* and re-embodiment of Maliseet language into Maliseet experience was the motivation for another immersive documentation project- *Wəlastəkwi (Maliseet) Cosmogenesis. Wəlastəkwi Cosmogenesis* is a Maliseet prayer of thanksgiving. It was first developed as *Nolosweltom* (I give thanks) by the native language teacher in the Maliseet language classroom as one of the language exercises. The prayer has a number of verses that express thanks for the sun, the moon, the stars, plants, animals, birds, etc. Each verse is accompanied by an illustration that I drew from the landscape on and around Tobique First Nation. The classroom prayer *Nolosweltom* antedates the graphic novel, whereas *Apotamkon* and *Wəlastəkwi Cosmogenesis* were developed at the same time. The strategy of surrounding the "reader" in Maliseet worlds is repeated in larger form in *Wəlastəkwi Cosmogenesis*. Each verse of the Maliseet prayer is represented by vertical panels measuring 36 inches by 70 inches. Each panel displays in graphic form the gratitude for the twelve relations,[14] a logographic of Tobique First Nation, a Maliseet/English text of the verse, a direction marker and personal signature, and an iconic Maliseet symbol etched in birchbark. The large panels are arrayed in a thirty-foot diameter circle aligned to the cardinal directions and appropriate interstitial points. The first line of the prayer faces east giving thanks to the sun.[15] As with the graphic novel, the background for each panel is the 360 degree landscape as viewed from the Tobique Rock. When a person steps into the circle s/he steps into Maliseet sacred space that invokes the voices of the ancestors, expands the experience of the present into the past as well as the future, and renews Maliseet cosmogony. Just like the novel *Apotomkon*, *Wəlastəkwi Cosmogenesis* transforms "reading" into the active *remembership* and re-embodiment of ancestral voices.

Both of these projects are still in progress. I would like to see *Apotomkon* transformed into a tablet device application that will allow the reader/viewer to be able to face east, hold the device in front of him/herself, and see the first page of the novel appear. As the viewer turns to her/his right, the next page will appear. Each page in turn will have the option of being read or having the story "told". My hope for *Wəlastəkwi Cosmogenesis* is a more permanent construction that will be located in Tobique First Nation, where community members can experience the Maliseet prayer. I plan the final installation to have audio playback for saying

---

14. The "reader" reads the first line of the prayer on the first panel while facing east. Each line of the prayer is then read in order going clockwise in order from east to east-south to south-east to south and so on. The relations depicted on the panels are- sun, moon, stars, grass, plants, trees, birds, animals, fish, earth, water, and wind.

15. The prayer begins with giving thanks to the sun while facing the sunrise. The Maliseet are one of the five Wabanaki nations, generally translated as "the people of the dawn".

and hearing each verse of the prayer. Both projects take advantage of contemporary genres of language use, as well as new technologies of "documentation" and language delivery, and they *remember* and re-embody Maliseet ancestral voices into contemporary Maliseet worlds. Rather than documentation as a practice preserving what is left of the Maliseet language before it disappears, I envision documentation as a means to bring prestige back to the Maliseet language through the use of new technologies, create new domains of language use, and provide catalysts for emergent forms of the Maliseet language in all their various vitalities.

4. Discussion: The future of indigenous languages

I began this paper with the proposition "If we are to save indigenous languages we need to remember ancestral voices." The complexities of that proposition included determining who "we" represents, what we mean by "save", how we define "indigenous" and "language", what "remember" entails, and who/what are "ancestral voices". The simple answer to who "we" represents is – "we" represents all stakeholders engaged in language documentation and revitalization efforts. But that answer is too simple. The roles of the stakeholders are equally complex. This paper acknowledges the significant contributions of language experts and the work they have done and continue to do on behalf of the communities whose languages are diagnosed as endangered, moribund, and extinct. My goal was to identify possible alternative documentary conceptualizations and practices that were initiated and implemented by indigenous language advocates/activists. Daryl Baldwin of the Miami Nation and many others working in relative obscurity represent the complex configuration of indigenous language advocates who represent their indigenous communities in awakening their languages. In my discussion of the Maliseet case, I described the work of Jeff Bear and the success he has had in broadcasting Maliseet across Canada. I am also a stakeholder who is both a native and an anthropologist with a particular set of skills that I bring to the task of "saving" the Maliseet language. What does "saving" endangered languages entail?

In the case of dead or extinct languages, it is less a matter of saving a dead language than it is a matter of awakening sleeping languages. For endangered or severely endangered languages such as Maliseet, the predictable interventions are documentation of the language as material artifacts such as lexicons, grammars, CDs, DVDs, etc. Less predictable are the creative projects that language advocates/activists initiate and develop themselves, be they translations of items from the dominant popular culture or new language practices in new domains of social interactions. These alternative documentary practices do not subscribe to

the salvage impulses of expert preservation interventions. They are the generative relations of emergent vitalities. In this context, how should we define "indigenous"?

There will be constant debates on what the term signifies, but for the purposes of this paper it designates the populations who have deep histories in particular landscapes whose languages have helped them negotiate symbiotic relations of sustainability and renewal. How, then, do we define "language" and what forms does *remembering* take?

It should be clear that I regard language as more than a linguistic code that is subject to *dismemberment* and analysis. I conceive of language as the broader semiotic code through which a community of speakers *remembers* all the social relations that make indigenous worlds vital and viable. *Remembering*, then, is not only the cognitive process of recall, but also the social process of inclusion and integration of all intersubjective relations, including the documents of language preservation as well as the ephemera of everyday communication. *Remembering* necessitates the re-embodiment of language, landscape, and spirituality as living Maliseet cosmogony.

Who and/or what are ancestral voices? Every language community has ancestral voices. Ancestral voices are never in the remote mythic past for indigenous communities, who have maintained intimate connections to their ancestral homelands. They continue to echo along the river valleys, the mountains, and across all landscapes as reminders of continuity and possibility. They are heard in every word, phrase, or expression uttered in indigenous languages. Documenting ancestral voices should not ossify those voices by preserving them in texts that sit unused on a bookshelf. Rather, we can produce documents that preserve the life of language as used between interlocutors but we must never trust the documents to "save" the language. We language experts, advocates/activists, community members are the ones who provide the crucial breath of language vitality.

I do not suggest that my approach for promoting emergent vitalities will work for all endangered indigenous languages. I do argue that some language situations may benefit from changing the metaphors from language death and extinction to sleeping, from endangered languages to emergent vitalities. There are a number of reasons I advocate changing our language intervention metaphors from language endangerment and subsequent language death to emergent vitality. Despite the relative obscurity of their work, community advocates/activists continue to rethink the metaphors used by language experts and breathe life back into their languages. For communities awakening their sleeping languages, the sleeping metaphor encourages creativity and emergent vitality. They are *remembering* ancestral voices. Not just in the cognitive sense but also in the cultural sense. The ancestral voices are becoming members of their descendent communities once again. Furthermore, switching metaphors of language death and extinction to

sleeping languages offers linguistic science a new prospect for research in emergent languages. Both linguists and indigenous language communities benefit from the shift in metaphors.

"Safe" languages, as defined by UNESCO, use information and communication technologies as part of their vitality and they serve as exemplars of adaptation, creativity, and expanding communicative domains. However, members of communities whose indigenous languages are considered "endangered" or "severely endangered" are exploiting the same kinds of ICT for adaptation, revitalization, and communicative creativity (Penfield et al. 2006; Dyson et al. 2007; Wilson & Stewart 2008). Native North American examples include inter-medial storytelling (Neely 2012), multi-register poetics (Webster 2012), curriculum and technology language instruction (Loether 2009), and data base archival and delivery systems (Berez & Holton 2006) to highlight some of the many interventions designed by community advocates/activists in partnership with linguists and other language professionals. Documentary linguistic practices also recognize and attempt to capture cross-modalities (Himmelmann 1998: 182–183) as part of the communicative event and current information and communication technologies make multi-modal documentation a flexible and invaluable tool. However, despite the benefits of technologically enhanced recording tools and archival and delivery systems, the primary goal is to document the communicative event in some permanent form that can be accessed by speakers, learners, and professionals. Michael and Dunn, working with the Badimaya People of Western Australia on developing their ICT cultural preservation project comment, "cultural preservation cannot be achieved by ICT alone: It requires the spiritual element behind the history to be actively reinvigorated into a community to make its presence felt in a long-lasting manner. Culture is something that is alive and ever-changing. In brief, it is not machinery that transforms society, repairs institutions, builds social networks or produces democratic culture; it is people who make things happen" (2007: 173). The collaboration between indigenous language advocates/activists and language professionals reflects how people "make things happen". I argue, then, a conceptual shift needs to take place in the various documentary practices away from documentation of communicative events to initiating communicative events where documentation can happen. Equally important, the conceptual shift away from language death to emergent vitalities creates possibilities for new forms of language use that utilize the skills of indigenous language advocates/activists. The variety of documentation materials and practices by both linguists and community language advocates need not be antithetical or mutually exclusive. Documentation as "alternative vitalities" is a critical step toward promoting emergent vitalities where the broad field of indigenous social relationships are *remembered*.

5. Conclusion. Back to the beginning

At a fundamental level all languages can be considered to be in emergent states as speakers creatively coin neologisms, expand linguistic domains, speak to one another on a daily basis, and most importantly, maintain and renew the intersubjective social relationships that make language an emergent vitality. This paper applies this perspective to languages that have been pronounced dead or extinct as well as those that are diagnosed as endangered. Rather than consider languages to be dead or extinct, we can better serve those indigenous language communities by changing the metaphor from extinct to sleeping languages. Doing so allows these emergent languages to offer linguistic science an exciting field for potential research. Among the benefits of "emergent vitalities" as an analytical and proactive stance toward language preservation and revitalization is the promise of collaborative work between language scholars and indigenous language advocates/activists. Finally, *remembering* the voices of the ancestors is an invocation of those voices by bringing the past into the present. Equally important, it is a renewal of ancestral voices as emergent vitalities of indigenous languages, cultures, identities, and spirituality. In this way, we language scholars, indigenous language advocates/activists, and community members participate in living indigenous cosmogonies.

The innovations in documentary linguistics for holistically recording communicative events in all their social, cultural, and communicative complexity are invaluable for both linguistic science and indigenous communities. But the products of such practices must not stop with documentation. Rather, their ontological potential as alternative vitalities will be key to awakening sleeping languages. Equally important, shifting the reason d'être of documentation away from record production and archiving toward an emphasis on the initiation and participation in developing social relationships where documentation "happens" can be the life giving catalyst for emergent vitalities of indigenous cosmogonies. To that end, my documentation projects promote emergent vitalities by *remembering* and re-embodying the Maliseet language in all their social relationships. *Remembering* and re-embodying ancestral voices can be the critical factor in shifting endangered indigenous languages toward vitality. The dictionary entry for the Maliseet culture hero Koluskap *dismembers* the hero from the Maliseet community by not conveying the contextually rich relationship Koluskap represents for the community. My graphic novel is a document designed to serve as a catalyst to *remember* Koluskap as an essential aspect of emergent vitalities of Maliseet language, culture, and identity. The Maliseet version is another step toward Maliseet *remembership* of Maliseet language with stories, landscape, and cosmogony (see Figure 2). But, this is not enough. My hope is that the story, the revitalization of ancestral Maliseet

Figure 2. *Remembering* Koluskap in the Maliseet language

voices, will promote new social relationships where the story is *remembered* in the community, language, and landscape far into the Maliseet future.

Dedication

I dedicate this paper to the memory of my older brother Wendell Samuel Perley who passed away July 9, 2012. Wendell was a Maliseet warrior in the truest sense of the word. He was an advocate for residential school survivors, an active participant in Maliseet language vitalities, and a gifted artist. He has been and always will be an inspiring example for me. My best friend, I miss you.

References

Austin, Peter. K. 2008. *Living, Endangered, and Lost: One Thousand Languages.* Berkeley CA: University of California Press.

Baldwin, Daryl & Olds, Julie. 2007. Miami Indian language and cultural research at Miami University. In *Beyond Red Power: American Indian Politics and Activism since 1900,* Daniel M. Cobb & Loretta Fowler (eds), 280–290. Santa Fe NM: School of Advanced Research.

Basso, Keith. 1996. *Wisdom Sits in Places: Landscape and Language among the Western Apache.* Albuquerque NM: University of New Mexico Press.

Bateson, Gregory. 2000[1972]. *Steps to an Ecology of Mind*: *Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology.* Chicago IL: University of Chicago Press.

Berez, Andrea & Holton, Gary. 2006. Finding the locus of best practice: Technology training in an Alaskan language community. In *Sustainable Data from Digital Fieldwork: The State of the Art*, Linda Barwick & Nicholas Thieberger (eds), 69–85. Sydney: Sydney University Press.

Bird, Steven & Simmons, Gary. 2003. Seven dimensions of portability for language documentation and description. *Language* 79(3): 557–582.

Boas, Franz. 1966[1940]. *Race, Language, and Culture*. New York NY: The Free Press.

Carroll, John B. 1970[1956]. *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf.* Cambridge MA: The MIT Press.

Crystal, David. 2000. *Language Death*. Cambridge: CUP.

Csató, Éva Á. & Nathan, David. 2003. Multimedia and the documentation of endangered languages. In *Language Documentation and Description*, Vol. 1. Peter K. Austin (ed.), 73–84. London: Hans Rausing Language Documentation Project.

de la Cadena, Marisol & Starn, Orin (eds). 2007. *Indigenous Experience Today*. Oxford: Berg.

Denzin, Norman, Lincoln, Yvonna S. & Tuhiwai Smith, Linda. 2008. *Handbook of Critical and Indigenous Methodologies*. Los Angeles CA: Sage.

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

Dorian, Nancy (ed.). 1989. *Investigating Obsolescence: Studies in Language Contraction and Death*. Cambridge: CUP.

Dyson, Laurel Evelyn, Hendriks, Max & Grant, Stephen. 2007. *Information Technology and Indigenous People*. Hershey NJ: Information Science Publishing.

Edwards, John. 1992. Sociopolitical aspects of language maintenance and loss: Towards a typology of minority language situations. In *Maintenance and Loss of Minority Languages* [Studies in Bilingualism 1], Willem Fase, Koen Jaspert & Sjaak Kroon (eds), 37–54. Amsterdam: John Benjamins.

Eisenlohr, Patrick. 2004. Language revitalization and new technologies: cultures of electronic mediation and the refiguring of communities. *Annual Review of Anthropology* 33: 21–45.

Enduring Voices. ⟨http://travel.nationalgeographic.com/travel/enduring-voices/⟩ (17 September 2012).

Erickson, Vincent O. 1978. Maliseet-Passamaquoddy. In *Handbook of North American Indians*, Vol. 15: *Northeast*, Bruce Trigger (ed.), 123–36. Washington DC: Smithsonian Institution.

Ethnologue: *Languages of the World.* ⟨http://www.ethnologue.com/web.asp⟩ (21 September 2012).

Evans, Nicholas. 2010. *Dying Words: Endangered Languages and What They Have to Tell us.* Malden MA: Wiley-Blackwell.

Francis, David A. & Leavitt, Robert M. 2008. *Passamaquoddy-Maliseet Dictionary: Peskotomuhkati Wolastoqewi Latuwewakon.* Orono ME: University of Maine Press.

Grenoble, Lenore A. 2008. Endangered languages. In *Living, Endangered, and Lost: One Thousand Language,*. Peter K. Austin (ed.), 216–235. Berkeley CA: University of California Press.

Grenoble, Lenore. A. & Whaley, Lindsay J. 1998. *Endangered Languages: Current Issues and Future Prospects*. Cambridge: CUP.

Grenoble, Lenore. A. & Whaley, Lindsay. J. 2006. *Saving Languages: An Introduction to Language Revitalization.* Cambridge: CUP.

Harrison, K. David. 2007. *When Languages Die: The Extinction of the World's Languages and the Erosion of Human Knowledge.* Oxford: OUP.

Harrison, K. David. 2010. *The Last Speakers: The Quest to Save the World's Most Endangered Languages.* Washington DC: National Geographic Society.

Himmelmann, Nikolaus P. 1998. Documentary and descriptive linguistics. *Linguistics* 36: 161–195.

Hinton, Leanne. 2001a. Language revitalization: An overview. In Hinton & Hale (eds), 3–18.
Hinton, Leanne. 2001b. Sleeping languages: Can they be awakened? In Hinton & Hale (eds), 413–17.
Hinton, Leanne & Hale, Ken. 2001. *The Green Book of Language Revitalization in Practice*. San Diego CA: Academic Press.
Ives, Edward D. 1964. Malecite and Passamaquoddy tales. *Northeast Folklore* 6. Orono ME: Northeast Folklore Society, Department of English, University of Maine.
Jewitt, Carey. 2011[2009]. *The Routledge Handbook of Multimodal Analysis*. New York NY: Routledge.
Krauss, Michael. 1992. The world's languages in crisis. *Language* 68: 4–10.
LeSourd, Philip S. 2007. *Tales from Maliseet Country: The Maliseet Texts of Karl. V. Teeter*. Lincoln NB: University of Nebraska Press.
Living Tongues Institute for Endangered Languages. ⟨http://www.livingtongues.org/⟩ (20 October 2012).
Loether, Christopher. 2009. Language revitalization and the manipulation of ideologies: A Shoshoni case study. In *Native American Language Ideologies: Beliefs, Practices, and Struggles in Indian Country*, Paul V. Kroskrity & Margaret C. Field (eds), 238–254. Tucson AZ: The University of Arizona Press.
Maaka, Roger C.A. & Andersen, Chris. 2006. *The Indigenous Experience: Global Perspectives*. Toronto: Canadian Scholars' Press.
Maaka, Roger & Fleras, Augie. 2005. *The Politics of Indigeneity: Challenging the State in Canada and Aotearoa New Zealand*. Dunedin: University of Otago Press.
Mandelbaum, David G. 1985[1949]. *Edward Sapir: Selected Writings in Language, Culture, and Personality.* Berkeley CA: University of California Press.
Miami Nation of Oklahoma. 2008. *myaamiaki eemamwiciki: Miami Awakening*. Upstream Productions.
Michael, Katina & Dunn, Leone. 2007. The use of information and communication technology for the preservation of Aboriginal culture: The Badimaya People of Western Australia. In *Information Technology and Indigenous People*, Laurel Evelyn Dyson, Max Hendriks & Stephen Grant (eds),174–174. Hershey: Information Science Publishing.
Minde, Henry. 2008. *Indigenous Peoples: Self-Determination, Knowledge, Indigeneity.* Delft: Eburon.
Moseley, C. 2010. *The Atlas of the World's Languages in Danger*, 3rd edn. Paris: UNESCO Publishing. ⟨http://www.unesco.org/new/en/culture/themes/endangered-languages/atlas-of-languages-in-danger/⟩ (28 February 2012).
Mühlhäusler, Peter. 1996. *Linguistic Ecology: Language Change and Linguistic Imperialism in the Pacific Region*. London: Routledge.
Mühlhäusler, Peter. 2003. *Language of Environment-Environment of Language: A Course in Ecolinguistics*. London: Battlebridge.
Neely, Amber A. 2012. Tales of tradition and stories of syncretism in Kiowa language revitalization. In *Telling Stories in the Face of Danger: Language Renewal in Native American Communities,* Paul V. Kroskrity (ed.), 90–114. Norman OK: University of Oklahoma Press.
Nettle, Daniel & Romaine, Suzanne. 2000. *Vanishing Voices: The Extinction of the World's Languages*. Oxford: OUP.
Niezen, Ronald. 2004. *A World beyond Difference: Cultural Identity in the Age of Globalization.* Malden MA: Blackwell.

Penfield, Susan D.; Cash Cash, Phillip; Kaleimamoowahinekapu Galla, Candace; Yetsastˆslahawi Kahˆtokt/Williams, Tracy & ShadowWalker, Depree. 2006. *Technology-Enhanced Language Revitalization.* Tucson AZ: University of Arizona Press.

Perley, Bernard C. 2002. Language, culture, and landscape: Preserving aboriginal deep time for tomorrow. Paper given at the UNESCO Conference Protecting the Cultural and Natural Heritage in the Western Hemisphere: Lessons from the Past; Looking to the Future, Harvard University, December 5–7. ⟨www.projects.gsd.harvard.edu/heritageintheamericas⟩

Perley, Bernard C. 2009. Contingencies of emergence: Planning Maliseet language ideologies. In *Native American Language Ideologies: Beliefs, Practices, and Struggles in Indian Country*, Paul V. Kroskrity & Margaret C. Field (eds), 255–70. Tucson AZ: University of Arizona Press.

Perley, Bernard C. 2011. *Defying Maliseet Language Death: Emergent Vitalities of Language, Culture, and Identity in Eastern Canada.* Lincoln NB: University of Nebraska Press.

Perley, Bernard C. 2012a. Zombie linguistics: Experts, endangered languages and the curse of undead voices. *Anthropological Forum* 22(2): 133–149.

Perley, Bernard C. 2012b. Last words, final thoughts: Collateral extinctions in Maliseet language death. In *The Anthropology of Extinction: Essays on Culture and Species Death*, Genese Marie Sodikoff (ed.), 127–42. Bloomington IN: Indiana University Press.

Perley, Bernard C. 2012c. Silence before the void: Language extinction, Maliseet storytelling, and the semiotics of survival. In *Telling Stories in the Face of Danger: Language Renewal in Native American Communities,* Paul V. Kroskrity (ed.), 184–204. Norman OK: University of Oklahoma Press.

Rymer, Russ. 2012. Vanishing voices. *National Geographic* 222(1): 60–93.

Schmidt, Annette. 1990. *The Loss of Australia's Aboriginal Language Heritage*. Institute Report Series. Canberra: Aboriginal Studies Press.

Shaw, Karena. 2008. *Indigeneity and Political Theory: Sovereignty and Limits of the Political.* New York NY: Routledge.

Smith, Linda Tuhiwai. 2005[1999]. *Decolonizing Methodologies: Research and Indigenous Peoples*. New York NY: Zed Books.

Steward, Julian H. 1972. *Theory of Culture Change: The Methodology of Multilinear Evolution*. Champaign IL: University of Illinois Press.

Stewart-Harawira, Makere. 2005. *The New Imperial Order: Indigenous Responses to Globalization*. New York NY: Zed Books.

Thornton, Thomas F. 2008. *Being and Place among the Tlingit*. Seattle WA: University of Washington Press.

Tsunoda, Tasaku. 2006. *Language Endangerment and Language Revitalization: An Introduction*. Berlin: Mouton de Gruyter.

Tuan, Yi-Fu. 1990[1974]. *Topophilia: A Study of Environmental Perception, Attitudes, and Values*. New York NY: Columbia University Press.

Tuan, Yi-Fu. 1977. *Space and Place: The Perspective of Experience*. Minneapolis MN: University of Minnesota Press.

Tuan, Yi-Fu. 1979. *Landscapes of Fear*. Minneapolis MN: University of Minnesota Press.

UNESCO. Language vitality. ⟨http://www.unesco.org/new/en/culture/themes/endangered-languages/language-vitality/⟩ (28 February 2012).

Urban rez.ca. ⟨http://www.urbanrez.ca/⟩ (21 September 2012).

Webster, Anthony K. 2012. To give an imagination to the listener: Replicating proper ways of speaking in and through contemporary Navajo poetry. In *Telling Stories in the Face of Danger: Language Renewal in Native American Communities,* Paul V. Kroskrity (ed.), 205–227. Norman OK: University of Oklahoma Press.

Wilson, Pamela & Stewart, Michelle (eds). 2008. *Global Indigenous Media: Cultures, Poetics, and Politics.* Durham NC: Duke University Press.

Index