DOCUMENTING ENDANGERED LANGUAGES AND MAINTAINING LANGUAGE DIVERSITY

Matthias Brenzinger
Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies, Japan

Tjeerd de Graaf
Slavic Research Center, Hokkaido University, Japan and Frisian Academy, the Netherlands

Keywords: endangered languages, ethnolinguistics, language documentation, language endangerment, language maintenance, language policy, language revitalization, language shift, mother tongue education.

Contents
1. Introduction
2. Language Endangerment and Endangered Languages
   2.1. Types and Extent of Language Endangerment
   2.2. Selected Case Studies of Endangered Languages
      2.2.1 /Xam, a Case of Physical and Cultural Genocide in South Africa
      2.2.2 Ainu in Japan
      2.2.3 Nivkh in the Russian Federation
      2.2.4 North Frisian in Germany
      2.2.5 Amazigh (Berber) Languages in Northern Africa
3. Language Documentation
   3.1. Why Study Endangered Languages?
   3.2. Documenting Endangered Languages
   3.3 Assessing Language Vitality and Endangerment
   3.4. Language Maintenance and Revitalization
4. Concluding Remarks
Acknowledgments
Glossary
Bibliography
Biographical Sketches

Summary

Several scholars predict that up to 90% of the world’s languages may well be replaced by dominant languages by the end of the twenty-first century, which would reduce the present number of almost 7000 languages to about 700. This review article attempts to describe processes that are underlying this severe threat to the majority of the languages currently spoken. However, the central focus of discussion will be on aspects related to the documentation and maintenance of the world’s linguistic diversity.

The main causes of language endangerment are presented here in a brief overview of the world’s language situation. Selected studies on endangered languages provide insight
into the case-specific aspects of these language shift situations. Finally, language documentation and maintenance efforts of scholars, international institutions, such as provided within the framework of UNESCO, as well as the role of academic cooperation in recent language documentation activities, will be discussed.

The fundamental task for linguists is research on the collection of data from endangered languages. Linguistics may preserve language data in documenting languages. However, it is the members of the speech communities that either uphold or give up languages. Only the speakers of endangered languages themselves can opt for and execute language maintenance or revitalization measures. Linguists and other scholars can assist communities in such attempts, for example by making language resources from archives available to them, by training community members to become language workers or even linguists, and also by helping to produce language learning and teaching materials. Linguists and community members should together take on the responsibility for documenting the wealth of linguistic diversity in order to pass on this legacy to future generations.

1. Introduction

In his article Who am I in this land? What people am I part of? Sergey Haruchi, representative of one of the minority peoples in the Russian North, gives the following account on his relation to the Nenets people and their language:

I experienced no shock on realising that I belonged to the Nenets people. Everything was formed during my childhood; both my father and my mother spoke Nenets. They did so without embarrassment, even in the presence of strangers. Regrettfully, though, the environment in which children now find themselves for the greater part of their time, at kindergarten, at school, has had an impact. They talk with their parents only in the mornings and the evenings. As a consequence, some children of indigenous intellectuals, not only from families of mixed marriages but even those whose fathers and mothers are both Nenets, do not know their native language. My own eldest daughter and son understand but do not speak Nenets, because they have no opportunity to practise it. I am not trying to use this as a reason to excuse myself. It is our fault and nobody else’s. Nevertheless, there are still children who know their native languages well and speak it fluently. When a cry of lamentation is raised that the language, the basis of the people’s culture, is sinking into oblivion and that books and textbooks should be published in native languages, I do not object to this; I even support it and assist to the best of my ability. But, after all, our parents did not teach us their native language by using books, we learned by hearing our mothers talk to us. It is first and foremost the mother who passes on the language, and a great deal depends on her because she spends more time with the children. It makes me wonder if it is right to put the blame for our children not speaking their native language on other people, for example, the Russian people, who allegedly impose their language on us. First and foremost, we must have strong desire to pass on the language. Our children ought to speak their native language.
This quote illustrates what can be considered to be the situation of minority languages in the Russian North, where, just as in many other parts of the world, the local languages are increasingly dominated and threatened by the official state languages, such as Russian in the case above.

At the 31st Session of the UNESCO General Conference (October 2001), the unanimously adopted *Universal Declaration on Cultural Diversity* recognized a relationship between biodiversity, cultural diversity, and linguistic diversity. UNESCO’s action plan recommends that Member States, in conjunction with speaker communities, undertake steps to ensure:

- sustaining the linguistic diversity of humanity and giving support to expression, creation, and dissemination of the greatest possible number of languages;
- encouraging linguistic diversity at all levels of education, wherever possible, and fostering the learning of several languages from the youngest age;
- incorporating, where appropriate, traditional pedagogies into the education process with a view to preserving and making full use of culturally-appropriate methods of communication and transmission of knowledge; and where permitted by speaker communities, encouraging universal access to information in the public domain through the global network, including promoting linguistic diversity in cyberspace.

2. Language Endangerment and Endangered Languages

A language is in danger when its speakers no longer pass it onto the next generation. Today, many speech communities of minority languages are shrinking and their languages will ultimately vanish, if these developments are not reversed. Children may no longer acquire languages even when they are still spoken by many thousands of elderly speakers.

2.1. Types and Extent of Language Endangerment

Language endangerment may be caused primarily by *external* forces such as military, economic, religious, cultural, or educational subjugation. It may also be caused by *internal* forces, such as a community’s negative attitude towards its own language or by a general decline of group identity. Internal pressures always derive from external factors. Together, they halt the intergenerational transmission of linguistic and cultural traditions. Many minority communities associate their disadvantaged social and economic position with their ancestral culture and language. They have come to believe that their languages are of no use anymore and not worth retaining. Speakers of minority languages abandon their languages and cultures in the hope of overcoming discrimination, to secure a livelihood and enhance social mobility for themselves and their children.

Endangered languages are not necessarily languages with few speakers. Even though small communities are more vulnerable to external threats, the size of a group does not always matter. The viability of a language is determined first and foremost by the general attitude of its speakers towards their heritage culture, of which their language may be considered the most important component. In this respect, the intergenerational
transmission of the language, i.e. teaching the children the heritage language, is the most important feature of language vitality.

The Suruaha, for instance, a small Indian community that lives in a remote area of Amazonia in Brazil, consists of approximately 150 members, and all of them—including the children—were monolingual in Suruaha at the time of their first contact with linguists. Despite the small size of the population, the community holds on to its language and traditional way of life in all domains. Natural or other disasters, such as epidemics or violent attacks, may threaten the physical survival of this small community, but leaving aside such external threats to the physical survival, the Suruaha culture and language seem to be “safe”. In contrast, many members of numerically large speech communities no longer pass their heritage languages onto the young generation. Such languages may still be spoken by thousands of elders, but nevertheless must obviously be considered as being endangered.

Language endangerment may arise when communities with different linguistic traditions live side by side. Such contacts involve an exchange of products as well as an exchange of cultural elements. Very often, the communities do not enjoy the same prestige in contact situations: a dominant vs. an inferior status may arise for specific reasons, such as numerical, economic, socio-historical or political strengths of each community. The communities with a lower status commonly acquire proficiency in the language of the dominant group. They may be inclined to relinquish their culture, including their language, and may decide to adopt the language and culture of the dominant community.

Statistical data related to language use may illustrate the extent of the problem of language endangerment. About 97% of the world’s people speak about 4% of the world’s languages; and conversely, about 96% of the world’s languages are spoken by about 3% of the world’s people. Approximately 85% of the almost 7000 languages of the world are spoken in only 22 countries. Some of these countries are home to large numbers of different languages: Papua New Guinea (almost 900 languages), Indonesia (up to 700), Nigeria (more than 500), India (almost 400), Cameroon (almost 300), Mexico (almost 250), Zaire (more than 200) and Brazil (more than 200). In these linguistically highly heterogeneous countries only a few languages have significant numbers of speakers and very few languages are assigned an official status within these states. Many of these languages are threatened by extinction. On a global scale, statistically less relevant are the threats to languages in linguistically more homogeneous parts of the world. In Europe, for instance, Sorbian is replaced by German, and the Saami languages in Sweden and Finland are threatened by the respective state languages, as Breton is in France.

The loss of speakers in one language is the gain of speakers of another language, except for cases of genocide. Languages are generally replaced when an entire speech community shifts to another language. Replacing languages are very often official state languages. For instance, on the African continent, several small communities replaced their heritage languages by Swahili in Tanzania, Somali in Somalia, or Arabic in the states of the Maghreb region. In other language shift situations, languages of wider communication have benefited, for example as trade languages or as a consequence of urbanization. Amharic in Ethiopia, Bambara in Mali, Hausa in Nigeria and Niger, and Wolof in Senegal are among those languages that have gained speakers at the expense of
smaller languages. On the African continent, languages of very small speech communities, however, are predominantly threatened by languages of other minorities. In Southern Ethiopia, for example, ‘Ongota is replaced by Ts’amakko (Tsamay), Kwegu (Koegu) by Mursi, Shabo by Majang and Harro by Bayso. Thus on the African continent the loss of speakers in one African language is typically the gain of speakers in another African language. English and French, the languages of the former colonial powers, are not replacing African languages, at least not for the time being.

In Europe most languages belong to the Indo-European language family, which represents the world’s largest group of languages with nearly 2 billion speakers. Within the Celtic subgroup Breton and Scottish Gaelic are endangered, while Irish Gaelic, despite the official support within the Irish Republic, is not safe. Within the Germanic subgroup, the East and North Frisian dialects in Germany are endangered as well as many dialects of Low Saxon (Nedersaksisch) in the Netherlands and Germany, the Mennonite Low German language ‘Plautdietsch’ and Yiddish. Rheto-Romance, the Romance language spoken in Switzerland is endangered. While Corsican, Occitan and Sardinian are still widely spoken, influence from the dominant official languages, i.e. French and Italian, respectively, is very strong. In the Western branch of the Slavic languages, Kashubian (Poland) and Sorbian (Germany) have only limited numbers of young speakers. The Indo-Iranian subgroup is represented in various countries by the Romani languages, which are endangered. Within the Uralic languages, the Saami languages are all endangered to various degrees. In the Scandinavian countries maintenance measures are undertaken for Inari Saami and Northern Saami. The Saami language forms in Russia, however, are almost extinct.

Language endangerment is a widespread phenomenon also in the northern part of Asia. The Altaic languages are spoken over a vast territory that extends from Turkey (Turkish) across Central Asia and Siberia to the Pacific Ocean (Yakut). Within the Turkic, Tungusic and Mongolic subgroups, most of the minor languages are endangered. In the northern regions of Yakutia in the Russian Federation, people who live in multi-ethnic communities experience pressure to abandon their heritage languages in favor of one of the two official languages, Russian and Yakut. As a result, many representatives of Evenki, Even (Tungus group), Chukchi and Yukagir (Paleo-Asiatic group) no longer speak their former languages.

According to Michael Krauss, only 20 of the 175 Native American languages in USA belong to the category of non-endangered languages, i.e. languages that are acquired by children informally at home. The extent of language disappearance in the Americas in pre-documented times has resulted in a large number of language isolates and distinct families with relatively few languages per family. Some scholars distinguish more than 100 language families and in addition leave a large number of individual languages unclassified.

The decline of native languages of North America was set off by the European conquests in the sixteenth and seventeenth centuries. Languages disappeared in large numbers because speech communities physically became the victims of warfare, new diseases, and forceful resettlement. Diseases, such as smallpox, were transmitted both unintentionally and intentionally; the deliberate transmission of some diseases was done with the clear
objective of genocide. Forceful resettlement also included slavery. The authorities set up language policies directed at eliminating heritage cultures and languages: native children were separated from their families in special boarding schools where the use of their language was forbidden. New value systems were forced upon the children and they were made to deny their own heritage and identity. This cultural and linguistic deprivation through formal education, that took place in many other colonial settings, such as in African countries, in Russia and in New Zealand, formed an essential part of colonial strategies to assimilate colonized peoples.

Native American languages have generally been replaced by the languages of the former colonial powers, i.e. Spanish, Portuguese, and English or local linguae francae. Colonial invaders considered multilingual settings as obstacles to evangelization and effective domination. Through this policy, Quechua and Tupí Guarani have gained speakers at the expense of smaller languages. After independence, indigenous languages by and large ceased to be taken into account within government policies.

In North America, at least 300 languages (estimates range from 300 to 600) are known to have been spoken by the indigenous populations. Many of these languages have become extinct, while many more are seriously endangered. Thus, Potawatomi and Onondaga have at most 100 speakers, and estimates place the numbers of speakers for languages such as Cayuga, Oneida, and Hare at 200 to 600. Salishan languages, though spoken in a large area in the northwest Pacific, are severely endangered. Wakashan constitutes a seriously endangered small family on Vancouver Island. Language maintenance programmes have been established for many of these languages.

Names of ethnic communities and place names are the only traces left of a large number of South American languages that have disappeared undocumented. In eastern Brazil and most of Argentina, as in many other parts of South America, the native population was eliminated and obviously their languages vanished in these genocides. The number of indigenous languages spoken in Brazil at the end of the fifteenth century, for example, has been estimated at approximately 1175. Efforts are being made to document the remaining approximately 180 languages.

In Australia over 250 languages were still spoken when Europeans arrived about 200 years ago. The majority of these languages have disappeared since then and speech communities of most of the remaining languages count fewer than 100 members. Bilingualism is widespread: seldom in other Australian languages, but predominately in English. With the exception of very few cases, the Australian languages are on the verge of extinction.

The extreme language diversity found in Oceania is illustrated well by the over 800 languages referred to as Papuan. The extent of language endangerment is not yet transparent, since less than 10% of the Papuan languages have been documented in any detail.

In the Andaman Islands of India a dozen languages form a distinct language family. The Northern and Central languages are already extinct, and the three severely endangered
languages of the *Southern Andamanese* group are spoken by between 50 and 250 speakers.

The *Sino-Tibetan* family includes Chinese, making it the second largest language family in the world in terms of total number of speakers. Less well-known is the fact that the *Tibeto-Burman* branch—with approximately 350 languages—constitutes 95% of all Sino-Tibetan languages. At least 20% of these languages are unclassified within the branch, and for many not even the numbers of speakers are known. Most of these languages are poorly documented and severely endangered. Extreme linguistic heterogeneity, together with a geographical spread over relatively inaccessible areas of northeastern India, Nepal, Bhutan, and Myanmar challenges attempts to present an overview of these languages.

### 2.2. Selected Case Studies of Endangered Languages

Any overview of the extent of language endangerment is incomplete. The main reason for this is a lack of information: there are territories where no language surveys have been conducted. Obviously, the situation of possible language endangerment is unknown in such regions.

Following this general overview, selected case studies will highlight some of the key issues central to language documentation and maintenance, such as the role of education, language policies, archives and records of endangered and extinct languages, and also the introduction of a writing system. The first case study deals with /Xam, an extinct language only known through language documents. As a second example, there is the Ainu language, with still a few speakers left. A few hundred, mainly older people, speak the Nivkh language, and for this language documentation and revitalization efforts are being undertaken. Two of the three Frisian languages are endangered, and language policies vary considerably in the Netherlands and Germany, where these languages are spoken. The last case study discusses aspects of the Amazigh language situation. Some of the Amazigh languages have disappeared without documentation, while others are on the verge of extinction and will most probably disappear. There is, however, hope for some of the Amazigh languages with a larger number of speakers to survive, as they receive official support in some of the Maghreb countries.

#### 2.2.1 /Xam, a Case of Physical and Cultural Genocide in South Africa

!Ke e: /xarra //ke is the motto of the state emblem of the new coat of arms, which was launched in post-apartheid South Africa in April 2000. This sentence from /Xam, a Southern Khoisan language spoken by one of the hunter-gatherer communities of Southern Africa, commonly referred to as San, is translated as “diverse people unite”. There are no speakers of /Xam left today. The phrase was coined on the basis of language documents, collected more than 90 years ago, and nobody knows, if native speakers of /Xam ever used these words in such a meaning. Nevertheless, the documents of the /Xam language form an important testimony of the autochthonous peoples of Southern Africa.

What happened to these people, their culture and language? A brief historical outline in answering this question follows Nigel Penn’s reconstruction. The colonial war against
the /Xam, which finally led to the extinction of /Xam identity and their language, is one of the cruelest events in African history. The ‘trekboers’ were the first colonialists to push into the interior of South Africa and to enter into a ruthless competition for natural resources. These frontier farmers appropriated the watering places for their cattle and killed large numbers of game with their guns. The San hunter-gatherer communities responded to this invasion into their land and to the massive destruction of their subsistence basis with guerrilla tactics. San attacked kraals at night, frequently killed the guards and stole the cattle or sheep. When possible, they drove the livestock into hidden places to consume the meat, or they simply killed the animals to destroy the farmers’ properties.

In the 1750s, the colonists decided to respond to the resistance with the aid of commandos. First, commandos were introduced to subjugate the Khoekhoe, pastoral communities speaking Central Khoisan languages, in the Western Cape. The Dutch intruders took their land and cattle and made the Khoekhoe unfree laborers. In the early missions against San, the commandos, which now consisted of Boers, Bastaards, Xhosa and Korana, had twin objectives: to crush the resistance and to capture women and children for use as laborers. The official policy at that time allowed the ‘trekboers’ to either chase San out of their areas or to force them into unfree labor. In 1774, a General Commando was set up to destroy any San resistance on a large scale. Usually, the various sections of the General Commando took off in surprise attacks at dawn, against individual kraals of sleeping San.

The numerically small San communities had no chance to strike back against guns and horse-mounted enemies. The number of casualties among members of the commandos was very small if any at all. Khoekhoe fighting in the first lines were killed or injured, but very few Boers were wounded or killed, while thousands of San males lost their lives. Instead of capturing them, the commandos usually shot San males, because they considered them as being of no economic value. More than 70 years after the General Commando, commandos in Bushmanland again killed large numbers of /Xam. The frontier farmers had given up attempts to ‘civilize’ the San and to change their ‘predatory existence’ into the life of pastoralists. The last remnants of the Cape San were hunted to extinction by the Boer and ‘Bastaard’ usurpers of Bushmanland in the 1870s. At that time, in addition to the commandos, also ‘hunting parties’ went out for the sole reason of killing San people. Thus, the war in the nineteenth century changed its objectives, as the farmers no longer aimed at merely breaking San resistance but completely eradicating San society.

The first words and some phrases of /Xam were recorded by Hinrich Lichtenstein during his travels between 1803 and 1806. In 1870, Wilhelm Bleek and Lucy Lloyd began to document the /Xam language. Their collection of extensive texts reveals that /Xam at that time was a widely spoken, viable language.

Bleek and Lloyd worked in the Cape Town prison with /Xam, who served sentences of imprisonment mainly for stock theft. In 1910-11, Dorothea Bleek visited remnant groups of /Xam, and found some were bilingual in /Xam and Afrikaans, while most others were already monolingual in Afrikaans. In the following decade, /Xam identity and the language disappeared altogether.
Bibliography

Advisory Committee on Endangered Language Research (2000). *Endangered Languages Research in the Netherlands*. A Report by the Netherlands Organization for Scientific Research NWO, The Hague. [This report has been prepared in the Netherlands with the aim to draw the attention of linguists and other specialists to the problem of endangered languages. After its publication a special program for the study of languages endangerment has been initiated].


Haruchi, S. (2002). Who am I in this land? What people am I part of? In: IWGIA Document *Towards a New Millennium, ten years of the indigenous movement in Russia*. Copenhagen, pp. 72-88. [Collection of documents and articles related to EL projects in the Russian Federation, in particular concerning the organization RAIPON of the Northern Peoples in Russia].


Biographical Sketches

Matthias Brenzinger is a tenured researcher at the Institut fuer Afrikanistik, University of Cologne, Germany and currently (2005/2006) visiting professor at the Research Institute for Languages and Cultures of Asia and Africa of the Tokyo University of Foreign Studies. He has organized several international conferences focusing on endangered languages and edited two volumes on language endangerment in Africa, i.e. Language Death. Factual and theoretical explorations with special reference to East Africa. Berlin, New York: Mouton de Gruyter, 1992 and Endangered Languages in Africa. Köln: Rüdiger Köppe, 1998. Fieldwork and publications on endangered languages on the African continent were conducted among the Yaaku in Kenya, Ma’a in Tanzania, Bayso in Ethiopia, Khwe in Namibia, Botswana, Angola and Zambia. He supported communities’ language maintenance activities, such as, the development of a practical orthography for Khwe and assisted as a consultant in the implementation of mother tongue education in Ethiopia. Since the mid-1990s he has been involved in UNESCO activities and projects related to endangered language issues, for example together with Akira Yamamoto as chairperson of the UNESCO Ad-hoc Expert Group on Endangered Language. He is a member of the standing committee of WOCAL (World Congress of African Linguistics).

Tjeerd de Graaf, associate professor of phonetics at Groningen University (the Netherlands) until 2003, has specialized in the phonetic aspects of Ethnolinguistics for the last 15 years. In 1990, he joined a Japanese expedition and conducted his first fieldwork with the minority peoples of Sakhalin. Since then he has contributed to various research projects on the endangered languages of Russia. In co-operation with colleagues in the Russian Federation and Japan, he focused on the use of sound archives for research on minority languages and cultures. Most of these research projects were funded by special grants from the European Union and the Netherlands Organization of Scientific Research NWO. Tjeerd de Graaf received a Doctorate Honoris Causa for his work from the University of St.Petersburg in 1998. Since 2002, he has been a board member of the Foundation for Endangered Languages (Great Britain). He is a research fellow at the Frisian Academy, which co-ordinates research on European minorities, in particular the language, history and culture of Frisian, one of the lesser used languages of Europe. In 2003, he spent a semester as visiting professor at the University of St. Petersburg and in 2004 and 2005, he was guest researcher at the Slavic Research Center of Hokkaido University, Japan.