

ENDANGERED LANGUAGE
RESEARCH
IN THE NETHERLANDS

AN OVERVIEW AND PROPOSALS

*Prepared
by the NWO Advisory Committee
on Endangered Language Research*

October 2000

CONTENTS

<i>Foreword</i>	1	
1 <i>Why study endangered languages?</i>	2	
1.1 Causes of language endangerment		2
1.2 Situations of language loss	2	
1.3 The value of the study of endangered languages		5
Linguistics and cognitive science		5
Language (pre)history and anthropology		5
The retention of cultural diversity	6	
Support for local initiatives	7	
1.4 Prioritizing the study of endangered languages		8
2 <i>Description and documentation</i>		10
2.1 Sources		10
2.2 Documentation and language maintenance / revitalization		11
3 <i>Endangered languages: an overview</i>		13
3.1 Endangered Languages in Europe and Northern Asia		13
Indo-European languages	14	
Uralic languages		14
Altaic languages		15
Languages of the Caucasus	15	
Paleo-Asiatic languages		15
3.2 Endangered Languages in Southern Asia and Oceania		16

Andamanese languages		16
Australian languages	16	
Austro-Asiatic languages		16
Austronesian languages		16
Dravidian languages	17	
Daic languages		17
Hmong-Mien languages		17
Indo-European languages	17	
Papuan languages	17	
Tibeto-Burman languages	18	
3.3 Endangered Languages in Africa		18
Afroasiatic	18	
Khoi-San languages	19	
Niger-Congo languages		19
Nilo-Saharan languages		20
3.4 Endangered Languages in the Americas		20
North America		21
Central America		22
South America		22
4 <i>Points of action</i>		24
4.1 International cooperation and coordination		24
4.2 Key areas in Dutch endangered language research		25
4.3 Creating the conditions for endangered language research		28
Funding		28
Support from the linguistic research community		29
Increased training facilities	29	
Better use of technological innovations		29
Provision for feedback to speech communities		30
5 <i>Further reading</i>		31
6 <i>Internet links</i>		33

Foreword

Around 85% of the estimated 6,000 languages of the world are spoken in only 22 countries. Some of these countries are home to more than 200 different languages: Papua New Guinea (850 languages), Indonesia (670), Nigeria (410), India (380), Cameroon (270), Australia (250), Mexico (240), Zaire (210) and Brazil (210). A large number of languages are threatened by extinction in these countries. Even in countries where the linguistic diversity is not as large, many languages are endangered.

This report, which provides an overview of the current situation of language endangerment and Dutch linguists' concern with that situation, was commissioned by the Netherlands Foundation for Scientific Research (NWO). Its purpose is to provide a basis for the evaluation of applications for funding submitted to NWO for endangered language research, and prepare the ground for international cooperation. NWO requested that we consider Dutch research traditions in the international context, with a view to ascertaining whether Dutch research has a specific contribution to make to endangered language research in particular regions or pertaining to particular language families, different from the contributions of research traditions which exist elsewhere. NWO further asked that we consider the extent to which new technology could contribute to the implementation of such research and the dissemination of its results.

We provide, first, a brief account of the main causes of language endangerment, and the compelling reasons for research and documentation of endangered languages in chapter 1. We consider different levels at which endangered languages can be documented, and the main factors which determine the type of documentation in chapter 2. In chapter 3, we take stock of the degree of language endangerment in different parts of the world, and the nature of the Dutch research tradition in these areas. The concluding chapter contains specific recommendations for points of action, and considers the particular strengths of Dutch research efforts in this area.

The committee would like to thank Bernard Comrie (MPI Leipzig), Cornelius Hasselblatt (Groningen University), Willem Adelaar, George van Driem, and Frederik Kortlandt (Leiden University) for their helpful comments, and Peter Meel and Monique van Donzel (NWO) for coordinating the committee and providing administrative support.

The NWO Advisory Committee on Endangered Language Research

Tjeerd de Graaf, Groningen University

Silvia Kouwenberg, NIAS/University of the West Indies

Maarten Mous, Leiden University

Pieter Muysken, Leiden University

Leo Wetzels, Free University Amsterdam

1 Why study endangered languages?

1.1 Causes of language endangerment

Endangered languages are not necessarily languages with few speakers; in fact, the size of a group hardly matters. The viability of a language is determined first and foremost by the general attitude of its speakers with respect to their traditional culture, of which their language is considered one of the most important exponents. The Suruaha, for instance, a small Indian people which lives in complete isolation in the Amazonian jungle of Brazil, consists of around 150 members, who all - including the children - speak Suruaha, and only Suruaha. Despite the small size of the population, the language and the traditional way of life are completely intact and nothing points to an imminent disappearance of the Suruaha culture. In contrast, some larger groups are not passing on their native languages to new generations.

Language endangerment arises in situations of contact between groups. Contact involves not only an exchange of cultural elements and products, but also of cultural prestige, which is often correlated with different degrees of technological advancement. A difference in technological know-how may lead to a sense of inferiority in the less highly developed group, which may then be inclined to relinquish its culture, including its language, in favour of that of the more highly developed group.

The question of whether the culture and language of a group should be regarded as an endangered one should be addressed in the light of the psychological, social and political circumstances that hold in a particular country. A reliable prediction as to how a people may fare when contact with a dominant culture is established can be derived from the way in which that dominant culture has dealt with minority groups in the past. In many cases, the attitude which national governments display towards their cultural minorities is not encouraging. Indigenous or aboriginal peoples are frequently regarded as backward and primitive, and - because relatively small minority groups often occupy large areas of potential commercial interest - as a barrier to economic progress.

1.2 Situations of language loss

The loss of one language is the gain of another: except in the case of genocide, a language is usually lost because speakers shift to another language. For instance, in several cases in Africa, groups have changed language allegiance to a non-colonial national language. This has strengthened Amharic (Ethiopia), Bambara (Mali), Somali (Somalia), and Swahili (Zaire/Congo). In other situations, a language of wider communication has benefited, often as a consequence of settlement in the urban centres. Thus, Arabic, Hausa, Jula, Lingala, Wolof, and other dominant regional languages have gained speakers at the expense of smaller languages. In a few cases, vocabulary loss has taken place without actual language death, resulting in a mixed language, as has

happened to several Bantu languages in Tanzania which have imported Swahili vocabulary. In sum, in Africa, language death is often the other side of the consolidation of other African languages and not of colonial languages, with Nigerian Pidgin English as a notable exception.

Similar cases can be found in the northern regions of the Russian Federation, where people living in multi-ethnic communities are under pressure to abandon minor languages, which has resulted in a dramatic decrease of the percentage of people speaking the language of their ethnic background. Thus, speakers of Evenki, Even (Tungus group), Chukchi and Yukagir (Paleo-Asiatic languages) in the north of the Republic Sakha (Yakutia) are shifting to one of the two official languages, Russian and Yakut. According to the results of the Russian census of 1989, only a third of the Yukagir ethnic group of circa 700 persons state Yukagir to be their mother tongue, Yakut and Russian accounting for the remaining two thirds. The situation is even worse than implied by these figures, as the number of speakers who know Yukagir well and use it fluently is not more than 20.

These situations contrast with that of the Americas, where it is mainly (former) colonial languages which benefit. Thus, speakers of native American languages have generally changed to Spanish, Portuguese, and English. In addition, we also find some instances where colonial rulers have tried to substitute a native *lingua franca* for a multitude of local tongues, which were considered obstacles to evangelization and effective domination. In this manner, Quechua, and TupíGuaraní have gained speakers at the expense of smaller languages. After independence, indigenous languages by and large ceased to be taken into account within government policy. With the possible exception of Paraguay, no South American nation has effectively derived part of its national identity from the use or even from the existence of a native language.

Language shift may be motivated by or concomitant with giving up a traditional economy that is central to the identity of the group, such as is the case for hunter-gatherers, cattle-nomads, or occupational “castes” such as smiths. In some African regions, conversion to Islam has played/plays a role in the spread of Hausa, Kanuri, Fulfulde and Arabic.

There are also cases of language death as a consequence of extinction of the people due to genocide, as is currently the fate of the people in the Nuba mountains in Sudan. Genocide is perhaps the most important single cause of language loss in North and South America. Genocide often takes place noiselessly, by spreading epidemics (either purposefully or unwittingly), forced relocation, or the destruction of the native habitat, which may lead to collective suicide, and at any rate to marginalization (alcoholism, prostitution, vagrancy) of the group at stake. In general, wars tend to have the effect of displacing and separating people with an adverse effect on language maintenance.

Whatever the cause, language loss leads to the irrevocable loss of human heritage. Whether people abandon their language in favour of a local language or in favour of one of the languages of (former) European colonizers, the death of a language affects linguistic diversity. In order to better realize both how threatening the danger of ongoing language loss really is, and how scarce the available material, Table 1 sketches the situation of the language endangerment status of the 27 languages distinguished in the *Chapacura* family, spoken in parts of Brazil and Bolivia. As we shall see in the following sections, the Chapacura family can be considered representative of many endangered and undescribed or underdescribed language families around the world.

1.3 The value of the study of endangered languages

There are a number of reasons why endangered languages should be studied and documented. These include scientific reasons, but also reasons concerning human rights issues and other motivations.

Linguistics and cognitive science

The ability to develop a language is regarded as one of the most characteristic and complex achievements in humans, not only by linguists, but also by psychologists and neurologists. The relation and interaction between the universal biological basis of language and the more general cognitive abilities which together enable us to acquire and use a specific language system (or grammar) is the subject of a discussion which is as controversial as it is interesting. To explain the phenomenon of language and to decide which part is grounded in biological predisposition and which in the general cognitive abilities and cultural representation systems in the human brain, it is obviously necessary to know how languages are structured and in which ways different languages construe the relation between form (sound shape, word form, sentence structure) and meaning. In other words, the question is what the ‘constants’ are in language, and what variations on these constant themes are possible.

For example, the highly endangered languages of Australia represent a pattern in which an intransitive subject (as in *Mary walks*) should be classed grammatically with an object of a transitive verb (as in *Charles sees Mary*) rather than with its subject (*Charles sees Mary*), as would be the case in most European languages. This phenomenon has fundamental consequences for our view of the clause and of basic grammatical functions, and would have gone unnoticed if these languages had not been documented.

Table 1 Language endangerment and documentation of the Chapacura languages (data reconstructed from de Lima Angenot 1997)

<i>Status</i>	<i>Language</i>	<i>Documentation</i>
---------------	-----------------	----------------------

Extinct or probably extinct	Jamará (Brazil)	no linguistic data available
	Kutiana (Brazil)	
	Tapoaya (Brazil)	
	Matáwa (Matáma) (Brazil)	
	Urunamakan (Brazil)	
	Kujuna (Brazil)	
	Herisobokono (Bolivia)	
	Itoreauhip (Bolivia)	
	Napeka (Nape) (Bolivia)	20-450 words
	Pawumwa-Wanham (Brazil)	
	Abitana-Wanham (Brazil)	
	Chapakura (Huachi, Tapakura) (Bolivia)	
	Kumana (Cautario) (Brazil)	
	Jarú (Brazil)	
	Uomo (Brazil)	
Miguelenho-Wanham (Brazil)	2,100 words, 16 lines of songs	
Muré (Bolivia)	translations of <i>Pater Noster</i> , <i>Ave Maria</i> , <i>Credo</i>	
Rokorona (Bolivia)		
1-10 speakers	Oro Wín (Brazil)	1,200 words and some sentences
	Kuyubi (Brazil)	2,500 words; fragments of a grammar
	Kusikia-Manasi (Bolivia)	no linguistic data available
20 speakers	Kitemoka (Kitemo) (Bolivia)	c360 words
100 -250 speakers	Kabishi-Waøam (Brazil)	83 words and 16 lines of songs
	Moré (Iten) (Bolivia)	3,800 words; 3,500 sentences; fragments of a grammar
	Urupá (Brazil)	c160 words
	Torá (Brazil)	387 words; 100 sentences

c1,000 speakers	Warí (Pakaás Novos) (Brazil)	Northern: 2,000 words, grammar fragments Southern: 3,000 words, one grammar
-----------------	------------------------------	--

Why different structural properties and categories appear in specific configurations is an exciting but extremely difficult question, which is the topic of various branches of linguistics. With the dramatic loss of languages, usually the least well-described ones, linguistics is faced with an enormous reduction of the empirical domain of its research area.

Language (pre)history and anthropology

The knowledge of indigenous languages is also the key to knowledge in several other disciplines, such as anthropology, archaeology, history and prehistory.

Language comparison and reconstruction can afford insights into the history and prehistory of the region where these languages are spoken. Sometimes relationships are discovered between languages and language families that are located at great distances from each other, giving us an understanding of the geographical origins of peoples and their possible migrations.

Similar information can be distilled from the conclusion that languages have been in contact with each other. Languages that exist side by side exchange words, subsystems or sentence constructions. In this way, so-called 'areal features' may originate, properties that are characteristic for specific regions, independent of the question whether the languages that are spoken there are genetically related. This kind of borrowing can give us a view of the time dimension of the contact situation. This is the case because the vocabulary of a language contains newer and older 'layers'. Depending on the question to which layer a word belongs, we can locate the borrowing nearer or further away in time.

Moreover, from the perspective of historical anthropology, it is of interest to determine, either within a language family or in a comparison of language families, which words occur in every language. This provides us with a proto-lexicon which informs us about the way of life and thinking of the original population. For instance, the names of plants may tell us whether these were nomadic (scion plants) or sedentary (seed plants) peoples. This approach can also provide clues as to the original location of a culture, especially when the lexical information can be combined with results from the (paleo)zoology and (paleo)botany of the area. The comparison and reconstruction of language thus adds to a better understanding of the (pre)history of indigenous languages and peoples.

The retention of cultural diversity

The value of the study of endangered languages can be considered against a background of the general value of preserving existing cultural diversity. Every culture represents an experiment in survival, of a unique and alternative way of life, of solving or evading problems. Loss of cultural diversity is therefore a loss of experience and knowledge that has proven its potential usefulness for mankind in general. Languages, besides being part of a nation's cultural heritage, constitute a complete and complex reflection of it. The loss of a language thus entails the loss of cultural heritage. The documentation of languages is thus also central to ethnobotany, ethno medicine, and the study of ritual and oral traditions.

Social value of documentation and language survival

The importance of language description has a clear ethical dimension, because the scientific interest can be seen as an expression of respect for an endangered civilization, which in turn gives it a sense of self-esteem. The ethical argument is highly relevant when the loss of linguistic and cultural diversity is the result of genocide. In a few cases, the recognition of the marginalised status of minorities has led to language maintenance programmes as part of policies to deal with social problems, as is the case in the United States. Often, aboriginal peoples live in poor or very poor countries, where the concept of minority status and related concepts of protection of minority rights and minority languages are not well developed.

Language loss leads to a situation in which a group which cannot express itself well in the majority language, comes to assume an inferior position in the margins of the dominant culture. Language documentation offers the possibility of setting up programmes for bilingual education, which will permit people to acquire literacy in their own language, fortifying group identity and self-esteem. The prestige which is conferred upon minority languages by standardization and the development of a native language curriculum rubs off on its speakers, a highly desirable result in situations where the dominant society and the minority groups themselves often share a negative view of these minority languages. Bilingual education will, furthermore, give minorities the means to better understand the society that wants to 'assimilate' them, and will allow them to defend themselves, especially involving matters which concern their own survival (land rights, exploitation rights to forests and other natural resources, complaints against pollution of drinking and fishing water, etc.).

Support for local initiatives

While national and local policies frequently state the desirability of the protection of minority rights, including minority languages, there is often a failure to implement such policies, sometimes through lack of resources, sometimes through lack of genuine interest. In such situations, support from outside interests for a policy of bilingual and intercultural education, and for

the development of language policy instruments is vital. Although scientists and politicians do not always agree on what exactly should be understood by integration, and, in particular, what role speakers themselves should play in determining the extent and the rate of their integration, all the parties concerned agree more and more that, at any event, bilingualism is advantageous for minority groups.

Some grassroots initiatives towards language maintenance have been taken, for instance, in Ecuador and Guatemala, where Indians make specific demands when their language is the subject of scientific inquiry. In exchange for information, they demand programmes for bilingual education, in which both the mother tongue as well as Spanish is taught. Minority groups also sometimes appeal to linguists to examine whether their language, which they no longer speak, was ever described in the past and, if so, to teach it anew. Such a request came, for instance, from the *Jenipapo Kanindê*, one of the indigenous groups in the north-east of Brazil who have become monolingually Portuguese. Unfortunately, nothing can be found on their language. This case shows that language loss comes to be lamented when a group loses its sense of cultural inferiority. It also shows how strongly minority groups identify language and identity, underscoring the social significance of language description.

The support which linguistic research provides for native speaker groups thus ranges from scientific attention, to language description, the development of instructional material, and the recovery of a lost language based on older materials (see §2.2).

1.4 Prioritizing the study of endangered languages

Not only is the extinction of a language an irreparable loss to linguistics, it is also an irrevocable loss to the cultural heritage of humankind. Knowledge of individual languages provides the empirical basis for linguistic theory and insight into how human thought is linguistically mediated. Documentation of individual languages is, thus, the key to understanding the human mind and human culture. Today, most languages are threatened with imminent extinction. Given the rate at which languages are currently lost, Krauss' pessimism seems warranted when he states that "the coming century will see either the death or doom of 90% of mankind's languages" (Krauss, in Hale et al. 1992:7).

On the basis of these considerations, the question of the priority of this type of research is moot. It is simply the case that it cannot be deferred to the future, as the languages that need to be documented will have disappeared altogether. Since human resources will be inadequate to document all endangered languages before they go extinct, strategic choices must be made. In principle it is of course desirable to make descriptions of every unique linguistic system, but faced with the practical limitations, choices must be made among the undescribed languages as to which deserve priority. Although it is often difficult to predict which languages have more to offer us, languages which

defy classification so far (*linguistic isolates*) are of special concern. This is so because the loss of a linguistic isolate means the loss of a language the properties of which cannot be reconstructed by research into related languages. Such languages are also the most likely to provide new evidence for language classification. Therefore, the uniqueness of the language insights that are likely to be lost with language death makes endangered isolates urgent candidates for research.

On every continent, there are also groups of languages which are endangered in their totality. Again, both from the point of view of the potential loss of data which may be considered vital for purposes of classification and the reconstruction of linguistic history, and from the point of view of the loss of unrecoverable typological data, such situations are extremely urgent. Recall also the fact that linguistic history provides insights into population movements and the history of contact between peoples.

Finally, the loss of contact situations and of special languages diminishes our potential insights into processes of language change and processes of register specialization. For this reason, the study of initiation languages and other ritual languages or registers, and the study of languages of occupational "castes", and of peoples dependent on other people in a client situation, is called for.

Based on considerations such as these, it may be possible to arrive at a prioritization of endangered language research. Additionally, the linguistic considerations may be supplemented by some of the other factors discussed in the preceding sections, such as the potential for language and culture maintenance by the group. Finally, it must be pointed out that areal expertise is located in particular universities in particular countries, and that international exchange and cooperation ensures that research of particular regions and languages does not duplicate research efforts elsewhere. The unique contribution of Dutch research to endangered language research and the expertise specifically available at Dutch institutions, thus forms a basis for prioritization at the national level, as will be explored in some depth in section 4.

2 Description and documentation

When we are talking about documenting endangered languages, a number of different types and levels of documentation can be thought of. Often, non-linguistic considerations determine the possibilities. Thus, the number of surviving speakers and their level of competence may impose severe limits on the work that may be carried out. Languages which may be regarded as 'very endangered' or moribund often cannot be fully studied, because no truly fluent speakers are available. This is the situation which often obtains when children have ceased to learn the language from their parents, so that the only living speakers are all adults, numbering perhaps not more than a handful of people. In such a case, extinction may only be a matter of years. In even more severe cases, limited insight into languages which are no longer spoken can be obtained from 'rememberers', people who are able to recall such a language as it was still spoken around them during their childhood years.

Other non-linguistic considerations include the availability of funding for equipment, the logistics of the fieldwork situation (accessibility), the familiarity of field workers with audio and audiovisual technology, etc.

One problem of research on endangered languages which deserves special mention, is that, often, a researcher needs to be familiar with the language that speakers are shifting to, as the only viable mode of access to the community. For instance, in some African situations, the necessity to know the language that is being adopted in some cases requires research on that language first, namely when it is insufficiently documented. This effectively burdens the researcher with the task of studying two languages rather than one, a situation which impacts on training, the length of research and on funding requirements.

2.1 Sources

For many extinct languages in the Old World the only sources of information are archeological findings, in the form of cuneiform tablets etc. These types of sources vary in their quality and size.

For many Amerindian languages we have catechisms or religious primers (*Doctrinas*), often dating back to the colonial era. If no other documentation is available, it is difficult to find out much about the language, because of limitations of the subject matter of these *Doctrinas*, which often take the form of literal translations from Spanish or Portuguese.

In the spirit of gathering encyclopaedic knowledge, characteristic of 18th century enlightenment, short word lists were compiled during that era. While these constitute a valuable source, spelling problems often impede good insights into the sound-systematic features of the languages in question, and of course, different word forms and sentences are not described in these sources.

In addition to these two types of sources, we find longer word lists and grammatical sketches, sometimes made by travellers or interested amateur

philologists who lived in the area for some time. These sketches contain useful material, although it has to be kept in mind that the material is frequently structured in ways that do not do justice to the unique properties of the language itself. It is an urgent task to survey and study the material to be found in various European archives, and make this material available on the internet.

It is only during the 19th and 20th centuries that more reliable and complete grammatical descriptions become available, often combined with dictionaries and text collections. These descriptive grammars are at present the main source for our knowledge of endangered and extinct languages. Much current work is a direct continuation of the tradition which originated in the 19th century, and produces grammars, dictionaries, as well as more theoretical analyses, often based on the results of lengthy periods of fieldwork.

The twentieth century has brought the technological innovation of sound carriers, complemented with video. We can document languages more completely by recording and filming speakers and interactions. The possibility of recording interactions has not just enriched our knowledge of grammar, morphology, lexicon, etc., it has also made it possible to systematically describe patterns of interactive language use.

Furthermore, information technology allows us to combine different types of information (texts, recordings, visual material) into integrated documentation systems. A major funding effort is needed to allow full use of these new techniques, however.

Finally, new technology can help make old materials accessible on new data carriers. Thus, sound archives based on different kinds of carriers, dating back to the early 20th century, exist for quite a few languages. Russian sound archives are a case in point. One of these consists of the recordings of a 1897 expedition among the peoples of Siberia, stored in the archives of the Pushkinsky Dom. It is one of the basic collections used by the INTAS project (carried out at Groningen University conjointly with institutions in St Petersburg), which, by copying onto modern sound carriers, aims to make these available for further research.

2.2 Documentation and language maintenance / revitalization

Currently, the documentation of endangered languages is faced with the challenge of making itself relevant to language maintenance efforts. Language descriptions are often couched in highly abstract language, making them unsuitable as tools in language maintenance, as requested by many communities where endangered languages are or were spoken. There is, therefore, the urgent question how we can bring the facts in the descriptions back to the speakers of the endangered languages. Given current technological possibilities and the explosion in available information, we need to organize data bases that give easy access to the data in different language descriptions.

Financial and logistic barriers rather than the limitations of linguistic knowledge impede progress in these areas.

Related to this is the problem of empowerment and control. Many existing descriptions are inaccurate and incomplete. Checking of language data by the communities of speakers that were the original source for these data is, therefore, desirable. Some language field workers succeed in training speakers as amateur linguists, thus contributing to the production of better descriptions as well as the viability of local efforts towards language maintenance. In co-operation with local linguists, a balance must be found between the attention paid to healthy languages that require research in order to aid their fuller development, and research on endangered languages.

Probably most language revival efforts are to be found in the United States. As an example, the *Breath of Life / Silent No More* workshops at the University of California in Berkeley, uses linguistic databases to help the native peoples of California study the language of their heritage. Of the approximately 100 different languages spoken by about 300,000 persons in California at the beginning of the 19th century, half have disappeared, and most of the other half are on the brink of extinction. There is not a single California Indian language that is learned by children as the primary language of their household. Even those people who know the language rarely use it: they may have nobody to talk to. The workshops assist groups of Indian people in reconstructing their ancestral language from publications, field notes, and audiotapes made by field workers in the past. Where elders still know the language, they help the workshop participants to learn how this language was used. Similarly, multimedia websites constructed for now extinct Australian aboriginal languages are frequently visited by descendants of the original speakers.

In this manner, material stored in sound archives, museum collections and libraries can be used to restore languages to the consciousness of tribal peoples, something of potentially great psychological and emotional importance, stimulating the well-being of such ethnic groups and providing the self-respect which comes with knowing one's heritage.

3 Endangered languages: an overview

Here, we will provide an overview of the extent of language endangerment within the language families represented in different parts of the world, and of the extent to which Dutch universities have research traditions of some permanence, either of particular language families or of languages of a particular geographical area, which may be considered endangered. Where possible, other major centres of research will also be identified. Where strong traditions of research by linguists and institutions from a particular country exist, it may be worthwhile to encourage that this research to be continued, while directing one's own attentions elsewhere. In this manner, the following overview will show the possibility for international collaboration in particular areas, as well as the unique contribution which Dutch research has the potential to make, as an indication of viable areas for future research.

Any overview of language endangerment will, of necessity, be incomplete. This is true in part because there are territories where no language surveys have been carried out; in such territories, the situation is insufficiently known to make reliable pronouncements on language endangerment. Moreover, because of the diversity and density of language stocks in many territories, it is almost an exercise in futility to try to say anything substantive about those regions in just a few pages.

While the gravity of a local situation of language endangerment is often a function of the number of languages spoken in that territory, several other factors may be considered in linking that situation to the urgency of the need for language research. At this point, it may be useful to recall the discussion in §1.4, where we indicated as particularly urgent (i) cases where language isolates are endangered; (ii) cases where an entire sub-family of languages is under threat of extinction, a situation which is all too frequently attested in different parts of the world; (iii) cases where special languages are endangered, including contact varieties, initiation languages or ritual languages, etc.

3.1 Endangered Languages in Europe and Northern Asia

The most important language families in this region are *Indo-European*, *Uralic* and *Altaic*. The languages of the Caucasus belong to various families and will be considered separately.

There is a long Dutch tradition in the study of these languages which goes back to the 17th century, when Nikolaas Witsen wrote his book on "Noord en Oost Tartarye", which is rich in detailed descriptions of the languages and cultures of the peoples of Russia, some of which have since disappeared. An international research team (Amsterdam, Groningen, Russia) is involved with the study of the data which it contains.

It is perhaps worth mentioning that many European countries are members of the Council of Europe and have ratified the European Charter for Regional or Minority Languages. This charter, which aims to protect the historical

regional or minority languages of Europe, took effect on March 1st, 1998. It states its purpose as follows:

"The aim of the Council of Europe is to achieve a greater unity between its members in order to promote their common heritage and ideals. Linguistic diversity is one of the most precious elements of this European cultural heritage. The cultural identity of Europe cannot be constructed on the basis of linguistic standardization. On the contrary, the protection and strengthening of its traditional regional and minority languages represents a contribution to the building of Europe, which, according to the ideals of the members of the Council of Europe, can be founded on genuine pluralist principles of democracy."

Indo-European languages

With nearly 2 billion speakers, the *Indo-European* language family is the world's largest group of languages. Within the *Celtic* subgroup, Breton and Scottish Gaelic are endangered, whereas Irish Gaelic has official support within the Irish Republic.

Within the *Germanic* subgroup, the East and North Frisian dialects (Germany) are endangered, whereas West Frisian (The Netherlands) is not. Also endangered are many dialects of Low Saxon (Nedersaksisch; The Netherlands and Germany), the Mennonite Low German language 'Plautdietsch' and Yiddish.

Within the *Romance* languages, Retho-Romanic (Switzerland) is endangered. While Corsican, Occitan and Sardinian still have many speakers, influence from the socially dominant official language (French or Italian) is very strong.

In the Western branch of the *Slavic* languages, Kashubian (Poland) and Sorbian (Germany) have only limited numbers of young speakers. Relevant research is carried out at Groningen University.

The *Indo-Iranian* subgroup is represented in various countries by the Romani language of the gypsies, a language which is seriously endangered.

Relevant projects are carried out at Leiden University, Groningen University (Plautdietsch, Yiddish, Sardinian, Frisian) and the Frisian Academy.

Uralic languages

Of the approximately 22 million speakers of *Uralic* languages, all but the Samoyed speak one of the Finno-Ugric languages. All *Samoyed* languages (Taymyr Peninsula, and along the lower course of the Yenisey River) are endangered – a situation where an entire subgroup is under threat. Of the *Finno-Permic* branch of Finno-Ugric, all languages spoken in the Russian Federation can be viewed as endangered. Least endangered are those languages which have official status next to Russian in different republics (Komi-Zyrian, Komi-Permyak, Mari, Mordvin, Udmurt). Very few speakers remain of the other Finnic languages, spoken in locations ranging from Karelia to Latvia. The *Saami* languages are similarly all endangered. In Scandinavia some measures are taken for the protection of Inari Saami and Northern Saami; the Saami language forms in Russia are all nearly extinct. The *Ugric* branch

includes several relatives of Hungarian in Western Siberia (Hanti, Mansi) which are endangered.

Dutch research projects of relevance are concentrated at Leiden and particularly Groningen Universities (INTAS projects, carried out in collaboration with Estonian and Russian colleagues).

Altaic languages

The Altaic languages are spoken over a vast expanse of territory extending from Turkey (Turkish), across Central Asia and Siberia, on to the Pacific Ocean (Yakut). Within the *Turkic*, *Tungusic* and *Mongolian* subgroups, most of the minor languages are endangered.

Research on Turkic and Tungusic languages is undertaken at Leiden University.

Languages of the Caucasus

This border area between Europe and Asia is famous for its multiplicity of languages, often with less than 1,000 speakers, and some of them strongly influenced by a dominant surrounding language, such as Georgian or Avar. Comrie (1981) distinguishes 38 languages within various subgroups. The status of some of these as separate languages is disputed, pointing to insufficient documentation to date. In addition, the Caucasian situation is of great interest as an area of extreme language contact.

München, London, Leiden are centres for the study of Caucasian languages.

Paleo-Asiatic languages

Many of the original languages of northeastern Asia are language isolates and not related to any language family or to each other. While some of these languages benefit from official recognition and/or local support programmes (e.g., the Ainu of Hokkaido, whose language is the only aboriginal language of Japan), most of these languages have no official status and are seriously endangered.

A large project has been set up by Japanese scholars for the description of these and other endangered languages around the Pacific. The University of Chiba is setting up fieldwork expeditions to Sakhalin and the Amur River to study the local languages and cultures, in particular Nivkh. Dutch involvement takes the form of collaborative efforts by the University of Groningen, which is involved in the Japanese projects and in projects with Russian research groups, and studies of Yukagir, Nivkh and other endangered languages at Leiden University.

3.2 Endangered Languages in Southern Asia and Oceania

The Dutch tradition of language documentation in Asia and Oceania dates back to the beginning of the 17th century, with some of the earliest documentation and grammars written in Dutch. The area is linguistically very rich, testifying to the antiquity of human habitation in the area. The typological and structural diversity of the area adds to the urgent need for research as well as language maintenance efforts.

Andamanese languages

Of this small family of about a dozen languages spoken in the Andaman Islands (India), the Northern and Central languages are already extinct. The *Southern Andamanese* group comprises 3 highly endangered languages (50-250 speakers).

Australian languages

Australia is - or rather was - home to over 250 languages, of which probably a quarter have already gone extinct, another quarter have fewer than 10 speakers, and the speaker numbers of most of the remaining languages is counted in tens rather than hundreds or thousands of speakers. Bilingualism is widespread, sometimes in other Australian languages, more often in English. In short, most of the Australian languages are irreversibly on the verge of extinction.

Relevant research is carried out in Australia, at the MPI (Nijmegen) and in Aarhus (Denmark).

Austro-Asiatic languages

Of over 200 *Austro-Asiatic* languages, spoken in an area extending from India to Vietnam, only a handful are currently not endangered. Documentation of most of the languages is sketchy.

Austro-Asiatic languages are studied in India (Mundan languages) and Europe (Mon-Khmer languages).

Austronesian languages

Austronesian consists of over a 1,000 languages, spoken from Indochina and Taiwan to Indonesia and New Guinea, to the island states of the Philippines, Micronesia, and beyond in the Pacific Ocean. Taiwan is the location of the *Formosan* subgroup, of which more than half the languages are extinct or nearly so, due to sinicization. The *Malayo-Polynesian* branch accounts for all the rest of territories. Some of these are endangered, whereas many are not, due to their geographic isolation.

In addition to research in Australia, Austronesian languages are studied in Germany (Frankfurt) and The Netherlands (Leiden, VU Amsterdam, MPI Nijmegen).

Dravidian languages

The *Dravidian* family comprises Tamil, Telugu, Kannada, and Malayalam, each of which has tens of millions of speakers. It also contains around 20 unclassified small languages, spread out over the Indian subcontinent, and another 60 languages, of which over half have fewer than 10,000 speakers, many only a few hundred speakers.

The Central Institute of Indian Languages carries out important research on many of the tribal languages of India.

Daic languages

Daic or *Kadai* consists of over 40 languages, spread over an area from India to Vietnam and into China. About half of these languages are endangered.

Hmong-Mien languages

The *Hmong-Mien* / *Miao-Yao* family consists of several dozen languages spoken in China, Vietnam, Thailand and Laos, many of which are threatened with extinction. The internal composition of the family and the number of member languages is actually uncertain, as classification of many languages is tentative, and some language labels may be cover terms for several languages.

Hmong-Mien languages are studied in the US.

Indo-European languages

With *Indo-European* being the best-studied language family in the world, there is still a surprising amount of work to be done on undescribed branches. Thus, all members of the poorly documented *Nuristani* family (mostly Afghanistan; several archaic languages) are threatened with extinction. A similar case is that of the over 20 undescribed *Dardic* languages (India, Pakistan, Afghanistan).

Of special concern are also the very many threatened unclassified *Indo-Aryan* languages. Thus, the *Rajasthani* branch (India) has more unclassified members than languages for which a solid classification has been made. We also mention Bote, Darai, Danuwar (Nepal), for which a *Dardic* affinity has been tentatively postulated, as insufficiently described and under threat.

Papuan languages

The extreme language diversity found in Oceania is well-illustrated by the over 700 languages referred to as *Papuan*. Less than 10% of these languages is documented in any detail.

Relevant research is concentrated in Australia, but is also carried out in Europe, including the Netherlands (VU Amsterdam, Leiden University, MPI Nijmegen).

Tibeto-Burman languages

The *Sino-Tibetan* family includes Chinese, making it the second largest language family in the world in terms of number of speakers. Less well known is the fact that at around 350 languages, the *Tibeto-Burman* branch constitutes 95% of the Sino-Tibetan languages, that at least 20% of languages within this branch are unclassified, and that for many languages not even the numbers of speakers are known. Most of these languages are thus both poorly documented and endangered. The situation is compounded by the combination of extreme linguistic heterogeneity with a geographical spread over relatively inaccessible areas of northeastern India, Nepal, Bhutan, and Myanmar.

Research is carried out at Leiden University on the languages of the Himalayas and Southern China. Research on the Tibeto-Burman languages is further carried out in China, India, Japan, the US and different European countries.

3.3 Endangered Languages in Africa

Research on African languages in the Netherlands is concentrated in but not limited to the research group of African Linguistics in the CNWS research school (Leiden University). Their policy is to diversify their research attention geographically and genetically in Africa. African countries with which Leiden has had intensive contacts during recent years include Burkina Faso, Cameroon, Congo, Ethiopia, Ghana, Morocco, Mozambique, Tanzania.

Several language isolates which are among the most endangered African languages deserve separate mention here: Shabo (600 speakers, Ethiopia), Ongota (8, Ethiopia), Hadza (200-400, Tanzania), Jalaa (30, Nigeria), Laal (300, Chad).

Afroasiatic

Berber languages, a group of 29 varieties, are threatened by political oppression. Some mixed Touareg varieties are of special interest for contact linguistics. Research is concentrated in Paris and Leiden.

A group of 192, *Chadic* languages are spoken in northern Nigeria, northern Cameroon, Chad, and Niger. Except for Hausa, all Chadic languages are endangered, many threatened by the spread of Hausa. European research is partly organized in a network called Mega-Chad.

Cushitic comprises 47 languages spoken in the Horn of Africa extending to Kenya and Tanzania. Several died out in recent years, while others are currently endangered. Research is carried out, i.a., in Addis Ababa, Cologne, London, Naples, and Leiden.

*Omoti*c constitutes a group of 28 languages, half of which have fewer than 10,000 speakers. We are still waiting for the first full reliable description of an Omotic language, but fortunately, research on languages of this group has been

growing since the opening up of Ethiopia, mainly carried out in Addis Ababa, London, and Leiden.

The *Semitic* subgroup, which consists of Arabic and its variants and Southern Semitic (Amharic, Tigre, Tigrinya, Gurage) in Ethiopia and Eritrea, is not generally endangered, although some individual languages in Ethiopia are.

Khoi-San languages

Khoi-San comprises a group of 35 languages spoken in southern Africa. All except Nama and Sandawe (in Tanzania) are endangered. A team in Cologne studies Khoi-San languages.

Niger-Congo languages

Atlantic is a group of 65 languages spoken along the coast from Senegal to Sierra Leone, of which 13 of fewer than 10,000 speakers.

Roughly 500 *Bantu* languages are spoken in the southern half of Africa. The group forms a unique research area because of the sheer numbers of closely related but different languages.

Benue-Congo languages excluding Bantu comprise an additional 200 languages, mainly spoken in Nigeria. Most languages are among the smaller and poorly described, many endangered, especially due to the expansion of Hausa.

Gur (100 languages spoken in and Mali and Burkina Faso and northern sections of Ivory Coast through to Benin) contains many small endangered languages.

Kru consists of 41 languages, mainly in Liberia and Ivory Coast. Of these, 8 languages have under 10,000 speakers.

Of special concern are *Adamawa* and *Ubangi* (86 and 71 languages spoken in the Central African Republic and surrounding countries) and *Mande* (10 out of 58 languages endangered). The classification of Mande is debated.

Ijo denotes a group of 10 languages spoken in south-east Nigeria, some of which are threatened; although assigned to Ijoid, the linguistic status of Defaka, which is moribund, is far from clear.

The *Kordofanian* subgroup (31 languages spoken in Southern Sudan) is endangered in its entirety, partly by genocide.

Kwa comprises 78 languages spoken from Ivory Coast to Benin.

Little research is carried out on Atlantic and Kru. Leiden University has had a tradition of research on some of the Benue-Congo languages. There are also some specialists in Germany. A research group in Bayreuth and Berlin is actively engaged with the Gur language. Hardly any research has been done on the endangered Kordofanian group of languages apart from research at Leiden University, and very recently at Paris. Relevant research further includes the Ijo languages (Port Harcourt, Nigeria), the Gbe and Akan clusters of Kwa (several research centres in Europe), the endangered Togo-Remnant clusters of the Kwa group (Leiden University), Ubangi (Paris; research concentrates on the largest languages), Adamawa (mainly by missionary linguists, some work has been

done at Leiden University). Research on Mande languages suffered from recent loss of leading specialists to other areas of research.

Nilo-Saharan languages

Central-Sudanic comprises 64, mostly small languages spoken mainly in Chad, Central African Republic, Congo-Kinshassa, Sudan.

Kuliak, a subgroup of which classification is debated, consists of 3 languages in Uganda, all of which are nearly extinct. This is a case where an entire subgroup stands to disappear.

Nilotic denotes a large group of 60 lects. Several languages are endangered.

Saharan languages, about eight in number, are spoken in the border area of Niger-Chad-Nigeria.

The remainder of the Nilo-Saharan languages are roughly in three areas: The Sudan-Ethiopia border area (Berta, Koma, Gumuz, Uduk, etc), Kunama, Nara in Eritrea, Songhai dialects at the Niger river, and languages or language groups such as Nubian, Fur, Maba, etc. in Sudan, Chad and the Niger-Nigeria border area.

Research on Central-Sudanic languages is carried out in Paris, Lyon, Brussels, Leiden. European research on Saharan languages is concentrated in Mainz and Vienna. Of the remaining languages, Songhai and the Nubian languages have attracted some attention from linguists, but languages such as Nara are virtually unstudied and its classification is extremely tentative.

3.4 Endangered Languages in the Americas

In the Americas, the survival of a few larger Amerindian languages, mainly in South America, contrasts with the disappearance of countless others, and the very serious endangerment of several more. The extent of language disappearance preceding any attempts at documentation is such that, today, America is distinct from comparable areas in the world by its very considerable number of language isolates, yielding an apparently unsurpassed genetic variety combined with a relatively small number of languages per family, where there is sufficient evidence to establish such families. Thus, the Amerind family of Greenberg's (1987) division of the Amerindian languages into three large families (Inuit, NaDene, and Amerind) is so large that it does not form a basis for description and classification. In some alternative classifications, well over 100 language families are distinguished, apart from a large number of individual languages left unclassified for lack of data.

In view of this situation, we discuss the endangered language situation of the Americas regionally rather than by language family.

North America

The native languages of North America have been in decline since the first Europeans arrived in the 16th and 17th centuries, with most of these languages either already extinct or currently in danger of disappearing. This language decline has resulted directly from the decrease of the population caused by the introduction of new diseases, warfare, forceful resettlement, etc. Another reason for this decline can be found in the language policies which were introduced by the authorities: native children were sent to special boarding schools where the use of their language was forbidden. The more tolerant policy adopted in recent years has come too late to halt this process for most of these languages.

In Canada, at least sixty languages are known to have been spoken by the original population, classified into *Eskimo-Aleut*, *Algonquian*, *Iroquoian*, *Siouan*, *Athabaskan*, *Salishan*, *Wakashan*, and *Tsimshian* languages, and the four isolates Beothuk, Tlingit, Kootenay and Haida. Many of these languages are extinct, while several are seriously endangered. Thus, Potawatomi and Onondaga have at most 100 speakers, while estimates place the numbers of speakers for each of the languages Cayuga, Oneida, Hare, Beaver, Kaska, Sekani, Kootenay, Okanagan, Shuswap, Thompson, Lillooet, Bella Coola, Sliammon Comox, Halkomelem, Haisla, Heiltsuk, Kwak'wala, Nootka, Coast Tsimshian and Haida at 200-600 speakers. Salishan languages, though spoken in a large area in the Pacific northwest, are highly endangered. Wakashan constitutes a fascinating but seriously endangered small family on Vancouver Island. For some of these languages teaching materials have been developed and are being used, although as yet with uncertain results.

According to Michael Krauss, only 20 of the 175 Indian languages in the United States belong to the category of non-endangered languages which are still learned in the traditional way: by children from their parents. As in Canada, the westernmost coastal areas of North America show the largest linguistic diversity with many non-related language families or isolates. This is the case in California where at the beginning of the 19th century more than 100 distinct Indian languages were spoken by over 300,000 people. Within one century the population was reduced to 10% and most of these languages became extinct or endangered. Some important examples of the surviving endangered languages of California include Cahuilla (50 speakers), Cocopa (400), Hupa (40), Karuk (40-60), Miwok languages (15), Pomo languages (75), Quenchan (400-500), Yokuts languages (70), Yurok (100). Special efforts are made to create a situation for language revitalisation (see 3.2). The methods used for this purpose provide an example of what could be done in other regions of the world where ethnic groups are expressing their wish to keep their own culture and to realize a revival of their language.

Research on North American Indian languages takes place at universities throughout Canada and the USA, with some individual research going on in

Europe as well. Particularly Algonquian, Salishan and Wakashan languages have been studied at Dutch institutions.

Central America

The situation in Mexico parallels that of Canada and the USA: a few native languages with relatively many speakers and a large number of families on the verge of extinction. The *Uto-Aztecan* family spreads from the US Southwest into large parts of Mexico. It includes Nahuatl, with 500,000 speakers, but countless extinct or seriously endangered languages as well. The *Otomanguean* and Mixtecan languages, to mention just a few of the other families, are very diverse, and some are poorly described.

Other smaller language families of Central America include the *Mayan* languages, which are widely spoken in the Yucatan peninsula and in Guatemala, and the *Chibchan* languages, which stretch from Costa Rica to Colombia.

Otomanguean and Mixtecan languages have been studied at Dutch institutions. Research centres exist also in the USA, in Mexico, in Paris and in Copenhagen.

South America

Many of the languages of South America have disappeared without leaving more than a tribal name and a few place-names behind. In huge areas, such as eastern Brazil and most of Argentina, the native population was eliminated and, as a consequence, the native languages as well. For instance, Rodrigues (1986) estimates that the number of indigenous languages spoken in Brazil at the end of the 15th century was around 1175. Today, there are about 180 languages left. Of the ones that are extinct, no complete description is available - not of a single one. Of the remaining languages, very few have been described in detail.

The total number of Indian languages in Argentina is at most 13, but many have either become extinct recently or are threatened with extinction.

In Brazil, the status of its 21 *Tupí-Guaraní* languages, altogether spoken by approximately 33,000 persons, ranges from endangered to moribund, with only three languages of more than 1,000 speakers. Languages of the *Macro-Jê* stock, which are scattered over a large part of the Brazilian territory, have around or fewer than 1,000 speakers. Smaller families include *Cariban* (21 languages, around 20,000 speakers), *Arawak* (17 languages, around 28,000 speakers), *Arawá* (7 languages, around 4,000 speakers), *Pano* (13 languages, around 5,000 speakers) *Tukano* (12 languages, around 7,000 speakers), several even smaller families, as well as a number of endangered linguistic isolates.

The number of Indian languages in Bolivia is approximately 35, including the two large highland languages Quechua (estimated at nearly 1,6 million speakers) and Aymara (estimated at well over a million speakers). However,

the total number of speakers of the other 33 languages is fewer than 100,000., and many have fewer than ten speakers.

The number of Indian languages still spoken in Chile today is probably 6 at most. Well known is the *Araucanian* language Mapuche.

The number of living languages in Colombia is somewhat higher than 60. While some languages have a fair number of speakers, the instability in the country contributes to the rapid breakdown of many groups.

Ecuador has 12 native languages if three *Tucanoan* groups are assigned separate linguistic status. The Indian population of the Ecuadorian Amazon region is estimated at between 30,000 and 60,000, and there are around one million speakers of Ecuadorian Quechua.

In the Guyanas (French Guiana, Suriname, and Guyana), over a dozen *Arawakan* and *Cariban* languages are spoken, all of which have small and declining speaker numbers. The Caribbean island representatives of these language families are extinct, with the exception of Garifuna (Black Carib), which was transplanted to Belize and thrives there.

There are 14 Indian languages in Paraguay. The number of tribal Indians in Paraguay may be estimated at 50,000. Several of the languages are seriously endangered.

The number of extant Indian languages in Peru is somewhere between 50 and 60, depending on which varieties are counted as separate languages.

There are around 38 native languages in Venezuela. This appears to be a reliable figure. Estimates of the Indian population of Venezuela vary between 50,000 and 145,000.

Of the language families in the Guyanas, particularly Cariban and Arawakan have been subject to Dutch research, and continue to be investigated at Dutch institutions, as well as at ORSTOM/CNRS in Cayenne. Given the relatively large number of completely undescribed linguistic isolates, Bolivia is currently an area which has priority in Dutch linguistic research. Descriptive research on Colombian languages is being carried out by Colombian and French linguists, with support of the French research organization CNRS. Venezuelan languages are studied in the Netherlands only to the extent that they are spoken in Guyana, Suriname, and Brazil. Important research centres for Brazilian languages are the Museu Goeldi in Belem, the State University of Campinas, Museu Nacional in Rio de Janeiro (Brazil), Rice University and the University of Oregon (USA), as well as the Netherlands (Leiden University, and VU Amsterdam).

4 Points of action

We perceive the responsibility for the documentation and preservation of linguistic diversity to be a global one. The notion of specific national responsibilities has undesirable post-colonial undertones. For sure, research in former colonies such as Indonesia and Suriname is called for because there are extensive research collections in the Netherlands with respect to these countries and some local networks are still in place. Likewise, there is good reason to plan research in other areas where strong earlier expertise has been gathered, expertise about languages of wider communication in the region, other neighbouring languages, political and social conditions, etc. Nonetheless, the vital importance of international coordination of research efforts is self-evident, and will be addressed in §4.1, whereas an overview of key areas for Dutch research is provided in §4.2.

We also feel that it is better to talk about research opportunities for Dutch researchers on endangered languages rather than about responsibilities: after all, documenting endangered languages is far from easy. The conditions under which this type of research is carried out often involve a certain level of hardship, and require considerable adaptability on the part of the researcher. The main prerequisites for successful research are considered in §4.3.

4.1 International cooperation and coordination

Given the scarcity of human and financial resources, it is important to contemplate international collaboration in the documentation of language diversity. It is not possible to develop hard-and-fast guidelines for such collaboration, since no single country has a consistent and transparent research policy in this area. Before countries were to agree on overall policies, most seriously endangered languages would have died out. However, this is not meant to imply that there is only chaos or, worse, wasteful overlap, in this domain.

First of all, UNESCO has established a clearing house for the documentation of endangered languages (ICHEL) in Tokyo in 1995. Many international meetings have taken place since then, either addressing the problem in general or on a regional basis (Africa, South America, the Russian Federation, etc.). At such meetings news is exchanged, international data bases are being developed, new priorities are set. Dutch descriptive linguists such as W. Adelaar (Leiden University) cooperate with professor S. Wurm (Research School for Pacific Studies, ANU, Canberra), coordinator of many activities in this field (Atlas of Endangered Languages, International Congress of Linguists, ICHL, etc.)

Second, there is a considerable amount of cooperation and coordination between the individual researchers at the grassroots level, simply because scholars want to come with a product that will be welcome and original, rather than duplicating someone else's efforts.

Third, there is a difference in the research traditions between individual countries. While Germany has often carried out larger-scale surveys and France has a tradition in ethnographically oriented lexicographical work, the Netherlands and the UK have a tradition of producing thorough grammatical descriptions of single languages. Thus, work by French and Dutch researchers in a single area would often already be complementary given the nature of the research itself.

Recently, the Max Planck Institute for Psycholinguistics (MPI, Nijmegen) has received a large grant from the Volkswagen Foundation to create software for all the projects on language documentation that will be funded by the Volkswagen foundation. It is to be hoped that this software will become the standard package for research in this field. Incidentally, the first set of specific grants awarded by the Volkswagen showed no regional concentration.

Turning specifically to the regional concentration of the research efforts of other countries, the following picture (grossly simplified) emerges:

Australia: Australia, the South Pacific and South East Asia

China: the languages of China

India: the Indian subcontinent

Japan: the Northern Pacific Rim

France: Francophone Africa and Caribbean, Colombia

Germany: global, with a focus on the Pacific, Africa

Russia: territories of the former Soviet Union

Scandinavia and Finland: global

United Kingdom: Ethiopia, South Asia

USA and Canada: global, with a focus on North America and the Caribbean

Among the countries of the so-called Third World, the amount of local expertise, and hence of potential for local collaboration, varies widely. For instance, Bolivia has virtually no linguists, while Brazil has very sophisticated research groups. Where feasible, western researchers should cooperate with linguists in the country where the research is carried out. Researchers should be encouraged to train local linguists where such a need exists, and to establish or join a local network in which the research can be continued, and language revival activities initiated.

4.2 Key areas in Dutch endangered language research

On the basis of (a) the amount of undocumented linguistic diversity; (b) research expertise in the Netherlands; and (c) the coordination of research efforts with other countries, several areas can be singled out as potential targets for research funding. These areas are listed below in alphabetical order.

Even so, we urge NWO not to take too static a view of these regions, not in the least because of changing political conditions. Research in Papua (Irian

Jaya), for instance, is both urgent and appropriate but at the time of writing not feasible due to political unrest. The Caucasus may be difficult at one time, relatively accessible at another. At present, research in Siberia is much more feasible than it was fifteen years ago. Ultimately, the strategic choice must be left to areal experts, based on both linguistic considerations, and other factors which impact on the feasibility of research.

Andean/Amazonian border. The region is characterized by a large number of linguistic isolates, considerable family-internal diversity, and large numbers of undescribed languages. It is a key area for understanding the linguistic prehistory of Latin America.

The southern part of the long border region of Brazil and Bolivia, Peru, Ecuador, and Colombia is currently being studied in several projects supervised by W. Adelaar (Leco, Arikapu), P. Muysken (Bolivian languages), and L. Wetzels (the Brazilian Nambikwara languages), but work further north is called for.

There is collaboration and coordination with research groups at the Universities of Stockholm, Lyon, Texas (Austin), with the Max Planck Institute at Leipzig, the Museu Goeldi in Belém (Brazil), and the German GTZ Centre EBI in Cochabamba (Bolivia).

Arctic Russia. The study of the languages in this area has been directed by F. Kortlandt (Leiden University/Spinoza research programme for languages of northern Eurasia and the Arctic region) and T. de Graaf (University of Groningen, collaborating with Russian scholars at the University of St.-Petersburg/NWO and INTAS projects), and also involves collaboration with the Japanese research programme on Endangered Languages of the North Pacific. Prof. Starostin has set up a data base of most of the Eurasian languages.

Research in this area is yielding new insights into the ancient relations between the peoples and languages of Eurasia, illuminating facets of their languages which are both theoretically and typologically as well as historically intriguing.

Among the most urgent and relevant examples in the Arctic zone of the Russian Federation are Nivkh, Kerek and Yukagir within the Paleo-Asiatic languages, Nenets, Enets and Nganasan within the Samoyed languages, and Orok, Even, Nanai and Dolgan within the Altaic languages.

Caucasus. Research in this linguistically extremely complex area was initially undertaken by A. Kuipers and is currently coordinated by R. Smeets. The emphasis is on northwest and northeast Caucasian languages, many of which are still undescribed. The languages Abkhaz, Circassian and Kabardian have been studied, as well as Hunzib. There is international collaboration with the MPI Leipzig (H. van de Berg, who works on Darginian).

Eastern Ghana. The Ghana Togo Mountain languages or Togo Remnant, a group of languages in the border area of eastern Ghana with Togo, are unusual in that they do not share the properties of the neighbouring languages, while they have a noun class system of the Bantu type. As a consequence their classification has been difficult. All fourteen of them count as highly endangered, including several on the brink of extinction due to pressure from the wider community languages, Ewe and Akan. Some other languages in this group died in the recent past. The communities are interested in language revival and, in general, in literacy programmes. In addition to F. Ameka's research on Likpe (Leiden University), research projects on Logba, Lolobi-Akpafu and Bowili are both desirable and feasible, with Dutch and Ghanaian candidates with knowledge of the wider community language available.

Guyanas. The Guyanas constitute a complex and linguistically highly diverse region where both traditional Amerindian languages are spoken and Creole languages as well as languages of Asian contract labourers. Dutch researchers have been active in the region for a long time, both with respect to Amerindian (B. Hoff and recently P. van Baarle and E. Carlin) and Creole languages (J. Voorhoeve, N. Smith, J. Arends, A. Bruyn). In addition, there is a strong tradition of anthropological and ethnohistorical research, and there are large research collections. French researchers are active in Cayenne, while a modest tradition of local research exists in (formerly British) Guyana. Given their complexity, the five Guyanas (Brazilian, French, Dutch, British, Venezuelan) constitute a promising prospect for international collaborative interdisciplinary research, ultimately perhaps with European community funding.

Himalaya. The greater Himalayan region is home to languages belonging to diverse families, including Tibeto-Burman, Austroasiatic, Andaman, Dravidian, and Daic, as well as several highly divergent language isolates. This rich genetic and typological diversity is threatened in particular by the politics of nation building. Dutch research in this area, coordinated by G. van Driem (Leiden University/Himalayan Languages Project), has involved collaboration with scholars and national governments in some of the pertinent areas.

The research has been very successful and yielded new insights into the genetic relationships of some of the languages in this area as well as the demographic history of the area, while also benefiting the local peoples, both through the prestige conferred on their languages by the research and the publications resulting from it, and by the implementation by local governments of recommendations to advance the role of these languages.

Indonesia. The Indonesian archipelago, including Papua (formerly known as Irian Jaya), is the home of some 1500 Austronesian and Non-Austronesian ('Papuan') languages. The linguistic diversity in this geographical area is

unparalleled, but to this date only some 5% of these languages have been described. In particular the languages spoken in Kalimantan, Sulawesi and the islands east of Lombok, including Papua, are in need of documentation, as many of these languages are threatened with extinction under the cultural, economical, and political dominance of the national language Bahasa Indonesia.

As a result of the historical ties between The Netherlands and Indonesia, the Dutch research collection on Indonesian languages is the world's largest by far. This rich scholarly heritage, and the availability of both research expertise and linguistic researchers (L. de Vries, H. Steinhauer, G. Reesink, C. Odé, M. van Staden, M. Klamer, A. van Engelenhoven, D. van Minde; W. van den Heuvel, N. Goudswaard), makes this region a prime focus area for Dutch research, in collaboration with Indonesian, Australian, or German (MPI has opened a field station in Jakarta) research groups. In fact, there is already extensive interational collaboration in this area.

Mexico, Guatemala, Belice. This area combines considerable ethnolinguistic diversity with a richly documented indigenous past and a highly turbulent present situation. Dutch researchers in the area (van Zantwijk, Jansen, and others) have been active in the ethnohistory and palaeography of the Chocho, the Mixtec, the Zapotec, and the Quiché-Maya among others. Language description has focussed on the Otomanguean language stock (J.W. Veerman-Leichsenring, E. Hekking, and, recently, M. Swanton). Other languages receiving attention are Mopan Maya (Belice) and Xinka (Guatemala). Research on the endangered languages in this area will be highly rewarding particularly because of the possibility of situating them in a socio-cultural context.

Pacific Northwest. This area in British Columbia contains a very great diversity of highly endangered Amerindian languages which have played a central role in linguistic debates and are also very interesting from the perspective of the prehistory of the Americas. In the 70's and 80's there was a strong research tradition in the Netherlands, directed particularly at the Salish and Wakashan language families (A. Kuipers, J. van Eijk, H. Nater, J. Rath, J. Timmers). Several linguistics students and young researchers (including G. Rowicka) envisage a revival of this tradition, extending research to the Na-Dene (Athabascan) language family.

Soudan and South West Ethiopia. The, roughly, thirty Kordofanian languages form a primary branch in Greenberg's and later classifications of Niger-Congo. Most Kordofanian languages are threatened by extinction as a result of the civil war in the Soudan. The urgency of research on this group of languages is very high. Although the war makes it impossible to do research in the area itself, it has also created a situation in which research among the refugees has become a sensible and efficient option, in Khartoum, Cairo, Nairobi and even, to some extent, in the Netherlands. The world expert in Kordofanian languages is T.Schadeberg (Leiden University) and he is in close contact with Dr. Nicholas Quint from CNRS in Paris who has recently initiated research on one of the

languages. The liberation movement in the Nuba mountains is very interested in the use of their languages in the schools that they have set up in the liberated areas, thus providing a context for language maintenance efforts.

South West Ethiopia is characterized by a high degree of diversity, including some unclassified languages. Also, the position of the Omotic language family within Afroasiatic is debated. Ethiopia has a language policy that supports the local languages and propagates education in the local languages. The country has invested considerable sums of money in linguistic research on the smaller languages. Despite this, many of these smaller languages are currently threatened because languages such as Amharic or Oromo impose themselves on the mixed communities to which people who formerly inhabited very remote regions are being relocated.

At present, Leiden University has three PhD projects on Cushitic-Omotic languages in this region, and cooperates with Addis Ababa, London and Naples in this area.

4.3 Creating the conditions for endangered language research

To accomplish the kind of research on and documentation of endangered languages that needs to be done now, a number of prerequisites have to be fulfilled:

- * **funding from NWO and other agencies**
- * **stronger and wider support from the linguistic research community**
- * **increased training facilities**
- * **better use of technological innovations**
- * **international cooperation**
- * **provision for feedback to speech communities**

Funding

In the past, research on endangered languages has frequently been funded by NWO on the basis of other, even extraneous considerations: theoretical interest, importance for language contact studies, historical or cultural interest. While we recognize the importance of these considerations, we urge NWO to acknowledge the documentation of the still extant language diversity as a goal of research in itself, and as a goal which in the next decade needs very strong financial support. This support is particularly needed since the documentation of language diversity cannot, by its nature, be a strong component of undergraduate teaching, the main source for university funding.

Support from the linguistic research community

Of the three research schools harbouring linguistic researchers in the Netherlands, currently only the Research School for Asian, African and Amerindian Studies (CNWS) acknowledges the study of endangered languages as part of its research programme. The other two schools, LOT and BCN, at best

allow such research in the margin of their activities, but do not explicitly endorse the documentation of language diversity as an important goal. While we understand the need for the differentiation in research goals between the different institutions, we feel that given its importance for ultimately understanding the nature of language, the documentation of language diversity should be on the agenda of all Dutch research institutions. Otherwise theoretical research programmes will only be parasitic on descriptive research, without contributing to it in a serious fashion.

Recently, five leading scholars in the CNWS proposed to establish an Institute for Comparative Linguistics within that research school (W.Adelaar, G.van Driem, F.Kortlandt, A.Lubotsky, T.Schadeberg, 1999). Within that institute, individual research projects would aim at language description and documentation in the field as well as reconstructing linguistic history and prehistory. While broader in scope than the present report, encompassing research on endangered languages as well as viable languages, that proposal was motivated by much the same concerns as those that form the basis for the recommendations put forward here.

Increased training facilities

There are few training courses at Dutch universities in linguistic fieldwork, a type of research requiring very specific expertise. Nor are there many fieldwork courses at universities in countries where endangered languages are spoken. There is, thus, a real need, both nationally and internationally, to set up new training programmes for young researchers, and also, we suggest, for established researchers who have so far not been engaged in descriptive work to do so.

Better use of technological innovations

Fieldwork training should include specific exploration of the ways in which new technologies for recording, storing and analysing data can be put to use in linguistic research, and training in the actual use of these technologies. Whereas linguists have generally not gone beyond sound recordings, new technologies can also handle a combination of sounds and images, and allow fast and cheap diffusion and sharing of research results, with speakers of endangered languages as well as with researchers in other countries.

Provision for feedback to speech communities

While most linguistic researchers acknowledge the importance of the documentation of endangered languages for language survival, the granting structure is such that feedback to the native speaker community which could support language survival efforts is generally not budgeted. In view of the desirability of permanence in the research efforts, which can best be achieved by the development of a local network which can continue the investigation, special demands may need to be made on the funding of endangered language research. Moreover, funding may need to be sought, possibly from other sources, to promote the recycling of research findings in text books, literacy materials, dictionaries, etc. As the internet gains a wider acceptance in non-industrial countries, it will become feasible and desirable to provide feedback to native communities using new technologies.

5 Further reading

- Adelaar, Willem 1998. The Endangerment of Native Languages in South America. In: Kazuto Matsumura (ed) *Studies in Endangered languages*. Tokyo: Hituzi Syobo, 1-15.
- Brenzinger, Matthias (ed) 1992. *Language Death: Theoretical and Factual Explorations with Special Reference to East Africa*. Berlin: Mouton.
- (ed) 1998. *Endangered languages in Africa*. Köln: Rüdiger Köppe.
- to appear. Proceedings of Round Table on Language Endangerment, Bad Godesberg, February 12-17, 2000.
- Comrie, Bernard 1981. *The Languages of the Soviet Union*. Cambridge: Cambridge University Press.
- Council of Europe 1995. *The situation of regional or minority languages in Europe*. Strasbourg.
- Crozier, D.H. and Roger M. Blench 1992 [2nd ed.]. *An index of Nigerian languages*. Dallas: SIL.
- Dieu, Michel and Patrick Renaud 1983. *Atlas linguistique de l'Afrique centrale (ALAC): Atlas linguistique du Cameroun (ALCAM). Situation linguistique en Afrique centrale. Inventaire préliminaire: le Cameroun*. Paris: ACCT; Yaoundé: CERDOTOLA/DGRST.
- Foley, William 1986. *The Papuan languages of New Guinea*. Cambridge: Cambridge University Press.
- Fox, Anthony 1995. *Linguistic Reconstruction: An introduction to theory and method*. Oxford: Oxford University Press.
- Grenoble, Leonore and Lindsey Whaley (eds) 1998. *Endangered Languages*. Cambridge: Cambridge University Press.
- Grupioni, Luís Donisete Benzi (ed) 1998. *Índios no Brasil*. São Paulo, Global Editora.
- Hale, K. et al. 1992. Endangered Languages. *Language* 68, 1-42.
- Hinton, Leanne 1994. *Flutes of Fire. Essays on California Indian Languages*. Heyday Books, Berkeley, California.
- Jazyki Mira, 1997. *Rossijskaja akademija nauk, Izdatel'stvo*. Moskva: INDRIK.

- Kasten, E. (ed) 1989. *Bicultural Education in the North*. Münster: Waxmann.
- Katzner, K. 1975. *The Languages of the World*. London: Routledge and Kegan Paul.
- Kaufman, Thomas 1990. Language history in South America. What we know and how to know more. In: Doris L. Payne (ed) *Amazonian Linguistics: Studies in Lowland South American languages*. Austin: University of Texas Press.
- Kortlandt, Frederik 1997. De buren van onze voorouders. *Spinoza 97*, Den Haag: NWO, 9-17.
- Krasnaja kniga jazykov narodov Rossii 1994. *Enciklopedičeskij slovar'-spravočnik*. Moskva: Academia.
- Krauss, Michael 1992. The world's languages in crisis. *Language 68*, 4-10.
- Ladefoged, Peter 1992. Another view of endangered languages *Language 68*, 809-811.
- de Lima-Angenot, Geralda 1997. *Fonotática e Fonologia do Lexema Proto Chapakura*. Unpublished M.A. Thesis, Universidade Federal de Rondônia, Guajará-Mirim.
- Loukotka, Chestmír 1968. *Classification of South-American Indian Languages*. Los Angeles: Latin American Center, University of California.
- Matsumura, Kazuto (ed) 1998. *Studies in Endangered Languages*. Papers from the International Symposium on Endangered Languages, Tokyo, November 18-20, 1995. Tokyo: Hituzi Syobo.
- Robins, R. and E. Uhlenbeck (eds) 1991. *Endangered Languages*. Oxford/New York: Berg.
- Rodrigues, Aryon Dall'Igna 1986. *Línguas Brasileiras*. São Paulo, Edições Loyola.
- Seifart, Frank 2000. *Grundfragen bei der Dokumentation bedrohter Sprachen*. *Arbeitspapier 36*. Institut für Sprachwissenschaft, Universität zu Köln.
- Smeets, Ineke and Willem Adelaar 1993. *Bedreigde Talen*. Department of descriptive and comparative linguistics, Leiden University. In K. Katzner 1975. *The Languages of the World*. London: Routledge and Kegan Paul.

- Steward, Julian. H. (ed.). 1946-59. *Handbook of South-American Indians* (7 vols).
- Tishkov, V.A. 1994. *Narody Rossii, enciklopedija. Naučnoe izdatel'stvo*. Moskva.
- Wurm, Stephen A. (ed) 1996. *Atlas of the world's languages in danger of disappearing*. Paris: Unesco.
- Wurm, Stephen A. and Shirô Hattori (eds) 1983. *Language atlas of the Pacific area*. Canberra: Australian Academy of Sciences in collaboration with the Japan Academy (*Pacific linguistics. Series C, Books ; no. 67*).
- Wurm, Stephen A. and Peter Mühlhäusler and Darrell T. Tyron (eds) 1996. *Atlas of languages of intercultural communication in the Pacific, Asia and the Americas*. Berlin [etc.]: Mouton de Gruyter.

6 Internet links

- Endangered Languages of the Pacific Rim
<http://www.elpr.bun.kyoto-u.ac.jp>
- Endangered languages list
endangered_languages_l@carmen.murdoch.edu.au
- Ethnic World Survey
<http://www.partal.com/ciemen/ethnic.html>
- Ethnologue - Languages of the World
<http://www.sil.org/ethnologue>
- Euromosaic Report
<http://www.uoc.es/euromosaic/>
- European Bureau for Lesser Used Languages
<http://www.eblul.org>
- European minority languages
http://www.smo.uhi.ac.uk/saoghal/mion-chanain/Failte_en.html
- Foundation of Endangered Languages

<nostler@chibcha.demon.co.uk>
 GeoNative
<http://cgi1.geocities.com/Athens/9479/welcome.html>
 Himalayan Languages Project
<http://iias.leidenuniv.nl/host/himalaya>
 Hokkaido University: Slavic research Center
<http://src-h.slav.hokudai.ac.jp>
 Indigenous Peoples of the World
<http://www.asahi-net.or.jp/~kv6y-frsw>
 Language futures Europe
<http://web.inter.nl.net/users/Paul.Treanor/enlang.html>
 Mercator-Education – European network for information, documentation and research into regional or minority languages in education
<http://www.fa.knaw.nl/mercator>
 MOST Clearing House Linguistic Rights
<http://www.unesco.org/most/In1.htm>
 Red Book of the Peoples of the Russian Empire
<http://www.eki.ee/books/redbook/>
 Research programmes supervised by Frederik Kortlandt
<http://www.let.rug.nl/~schaeken/kortlandt.html>
 Society for Endangered Languages/Gesellschaft für bedrohte Sprachen
http://www.uni-koeln.de/GbS/e_index.html
 Terralingua
<http://cougar.ucdavis.edu/nast/terralin/home.html>
 Teaching Indigenous Languages
<http://jan.ucc.nau.edu/~jar/TIL.html>
 The International Clearing House for Endangered Languages
<http://www.tooyoo.l.u-tokyo.ac.jp/ichel.html>
 The Languages of Russia – A Multimedia Website
<http://www.let.rug.nl/~bergmann/russia.htm>
 The Linguistlist Endangered Languages Homepage
<http://www.linguistlist.org/el-page>
 UCLA Language Materials Project: Learning Resources for less commonly taught languages of the world
www.lmp.ucla.edu/default.htm
 UNESCO Red Book on Endangered Languages
<http://www.helsinki.fi/~tasalmin/endangered>
 Unrepresented Nations and Peoples Organisation (UNPO)
<http://unpo.org>
 Miscellaneous
<http://cougar.ucdavis.edu/nas/terralin/elbiblio.html#anchor77337>
<http://coombs.anu.edu.au/coombspapers/otherarchives/cipsh-unesco-archives/>
<http://www.ammerlaan.demon.nl/LOSS.HTM#research>