

the 24th June 2000.

- Tsunoda, Tasaku. 1981. *The Djaru language of Kimberley, Western Australia*. Canberra: Pacific Linguistics.
- . 1998. *Fiirudo waaku to ippan gengogaku* ('Field work and general linguistics'). Pp.20-25 of the preprints of the 117th meeting of Linguistic Society of Japan, held at the University of Yamaguchi, on the 31st October to the 1st November, 1998. Kyoto: Linguistic Society of Japan.
- . 2000. *Fiirudowaakaa no rinri to ninmu* ('Role and ethics of field workers'). Gengo, Vol.29, No.2:106-12. Tokyo: Taishukan.
- . 2001a. Role and ethics of researchers and method of documentation. In Sakiyama (ed.) 2001: 261-67.
- . 2001b. Introduction. In Tasaku Tsunoda (ed.), *Basic materials in minority languages*, 1-12. (*Endangered Languages of the Pacific Rim*, 2001, B001), 261-67. (For distribution, contact elpr@utc.osaka-gu.ac.jp.)
- Wilkins, David. 1992. Linguistic research under Aboriginal control: a personal account of fieldwork in Central Australia. *Australian Journal of Linguistics*, Vol.12, No.1:171-200.
- Yamamoto, Akira. 1998. Linguists and endangered language communities[:] issues and approaches. In Matsumura (ed.) 1998:213-52.

The Holistic Approach to Endangered Language Documentation: Research Philosophy of the Himalayan Languages Project

George van Driem
Leiden University

The greater Himalayan region is the area of the greatest phylogenetic heterogeneity in Eurasia. Because of socio-economic and demographic pressures, this region is also the area with the highest density of endangered languages in Eurasia. The Himalayan Languages Project is an ongoing programme of Leiden University with a multinational research team of young investigators who venture to chart the indigenous languages of the Himalayas. These languages hold the key to the prehistory and cultures of the Himalayas and enable the interpretation of the archaeological record and the tracing of ancient population movements throughout the Asian heartland. Detailed holistic descriptions of languages render accessible typological data on fascinating linguistic features of this region and directly benefit the local people. The work of the research team of Ph.D. students and post-doc researchers of the Himalayan Languages Project yields a large body of detailed and diverse knowledge on the vanishing languages and cultures of the Himalayas. What lessons can we learn from the experiences of the Himalayan Languages Project?

First fruits of the research programme

The Himalayan Languages Project has received financial support from a dozen sponsors, but the single most significant contribution is the one and a half million guilder grant bestowed by the Netherlands Organisation for the Advancement of Scientific Research (NWO) in April 1993. The original goals of the research programme, specified in the original proposal and the 1st Interim Report of February 1996, were to produce the following:

- (1) 'five in-depth grammars of hitherto undescribed Tibeto-Burman languages in the Himalayas', each produced individually by a member of the research team under the guidance of the project director. Each grammar was to include a dictionary and analysed native texts;
- (2) spin-off in the form of articles on Tibeto-Burman historical comparison;
- (3) detailed knowledge of the indigenous cultures of the Himalayas, preserved for posterity;
- (4) spin-off in the form of typological and theoretical studies devoted to newly attested linguistic phenomena of major typological and grammatical interest in the languages described;

(5) 'a major new synthesis' on languages of the Kiranti group, with contributions by the project director, research team members and specialists in the field outside of the project.

Today we can look back and see that the research team of the Himalayan Languages Project has met and greatly exceeded these goals and shown itself to be a sustainable research programme. The main results thus far include:

- (1) Not just the originally envisaged 5, but a total of 11 comprehensive grammars and grammatical sketches have been completed to date.
- (2) Moreover, an additional 15 comprehensive grammars and grammatical sketches are currently being completed by the Himalayan Languages Project research team.
- (3) In addition, dozens of minor studies have been published on 15 different languages and language communities of the greater Himalayan region.
- (4) Extra funds were raised from sources as diverse as the Swiss watchmaker Rolex and donors who wish to remain anonymous. The allure of the research programme has drawn foreign researchers who in some cases brought in their own funding.

In addition to its main task of providing detailed documentation of endangered languages in the form of holistic language descriptions, other major results of the Himalayan Languages Project include the following:

- (1) the discovery of two languages previously unknown to science, each constituting a distinct subgroup unto itself within the Tibeto-Burman language family, i.e. Gongduk and Black Mountain in Bhutan;
- (2) the rediscovery in central Nepal of the last surviving Baram language community, a tongue which had been presumed extinct since the 19th century;
- (3) the exhaustive identification of all Tibeto-Burman language communities and subgroups in the 1990s;
- (4) the bringing about of a major paradigm shift in the field, leading to the abandonment of the obsolete Indo-Chinese or Sino-Tibetan model, the revival of the older, original Tibeto-Burman phylogenetic model, and the identification of evidence for several new higher-order subgrouping proposals within the language family;
- (5) the identification of the Karasuk language family on the basis of grammatical evidence for genetic relationship between Burushaski and the Yenisseian languages;
- (6) a two-volume compendium on the languages, language families, language communities and prehistory of the Asian heartland, entitled *Languages of the Himalayas: An Ethnolinguistic Handbook of the Greater Himalayan Region*,

complete with one hundred first-ever maps of many language areas in the Asian heartland;

- (7) reconstructions of population movements which led to the peopling of the Himalayas and dispersal of language families on the basis of careful correlation of the data on genes, languages and archaeology;
- (8) exposing and resolving the so-called Bangani Enigma of the western Himalayas, an elaborate hoax perpetrated on the international scholarly community;
- (9) designing an orthographically streamlined phonological transcription for Dzongkha which was adopted by the Royal Government of Bhutan as the official system of romanisation for the national language;
- (10) conducting the First Linguistic Survey of Bhutan at the behest of the Royal Government of Bhutan;
- (11) setting up and participating in the permanent Linguistic Survey of Bhutan at the request of the Royal Government of Bhutan, involving the documentation of the three most endangered languages of the kingdom;
- (12) a beneficial impact on the formulation of the official language policy of post-revolutionary Nepal;
- (13) helping to shape and formulate the language policy of Bhutan in close consultation with the relevant Bhutanese authorities;
- (14) setting up a comparative and historical Tibeto-Burman database on the internet, easily accessible and regularly updated to incorporate new data and novel insights into Tibeto-Burman historical phonology; and
- (15) editing a state-of-the-art anthology to be published by Mouton de Gruyter in Berlin.

Activities of the Himalayan Languages Project research team show a broad commitment to the field and the international scholarly community. The research team hosted the 1st Himalayan Languages Symposium in 1995 in Leiden and the 2nd Himalayan Languages Symposium in 1996 in Noordwijkerhout. Since then, the Himalayan Languages Symposium has become a regularly convening open scholarly forum, which has thus far convened in Santa Barbara (California), Pune (India), Kathmandu (Nepal), Milwaukee (Wisconsin), Uppsala (Sweden). The 8th Himalayan Languages Symposium in 2002 will be held at Bern University in Switzerland. The Himalayan Languages Project at Leiden maintains the permanent secretariat of the annually convening symposium. The Himalayan Languages Project also hosted the Old Chinese Seminar in 1995, the international Old Chinese Forum in 1996, and the 29th International Conference on Sino-Tibetan Languages and Linguistics in 1996. More information is available on all these matters in our 2nd Interim Report and at our website.

The research philosophy

But the main aim, task and output of the Himalayan Languages Project is holistic language documentation in the form of comprehensive grammars like the grammar of Yamphu, written by Roland Rutgers. Yamphu is an endangered language of the Himalayas spoken by a small community in the mountains of eastern Nepal. The language is rapidly dying out, and the Yamphu community is quickly assimilating to the national language Nepali. The grammar contains a detailed description and analysis of the phonetics and phonology, a detailed account of the morphophonological regularities of the language, and a description and analysis of the morphology, including a detailed account of the formally complex biactantial agreement system of the verb and a detailed characterisation of the meanings and operation of the typologically interesting aspectual, temporal and epistemic categories of Yamphu grammar. This book also contains an account of the syntax and morphology of all types of simple and complex clauses, including all sorts of subordination, indirect speech patterns and constructions with gerunds. The book includes illustrative conjugational tables of all of the various verbal paradigms found in the language and a bilingual Yamphu-English glossary, which also provides the necessary grammatical information. The book also contains a morphologically analysed native Yamphu text corpus, complete with interlinear glosses and an integral English translation. Finally, there is also an account of Nepali loans in the language and colour photographs of speakers and the language community.

I put before you today that the production of grammars of this type is the very best way to invest public and private funding that has been earmarked for endangered languages. The reasons for this are threefold:

(1) Such holistic language documentation is of the maximum benefit to the particular language community concerned.

The Limbu grammar produced by the Himalayan Languages Project has been used by the Limbu community as a model and a public resource for the production of language learning materials for use in primary education in the mother tongue in schools in eastern Nepal. The Limbu grammar is now being used by the Royal Nepal Academy in the preparation of a large Limbu-Nepali-English dictionary, particularly in view of the need to know which type of grammatical information must be specified for each entry in the dictionary, such as the conjugation of a verb and the phonological shape of its various stems, something which cannot be deduced from the infinitive form under which a verb is listed.

The grammar of Dzongkha produced by the Himalayan Languages Project is being used in the same way by the Royal Government of Bhutan in the production of language learning materials for primary education in the mother tongue in schools in the Kingdom of Bhutan, but also in the preparation of dictionaries and computer interfaces. The official romanisation of the national language was also adopted from the phonology included in this grammar.

Even in cases where the process of language death is irreversible and the extinction of the language cannot be stopped, holistic grammars with text corpus and glossary provide the most valuable contribution to the language community. The members of the Dumi ethnic community in Nepal cherish the detailed and comprehensive Dumi grammar produced by the Himalayan Languages Project because it contains the only account, not only of the nearly extinct language of their ancestors, but also of the native lore, myths, eschatology, pantheon and legends of their people.

The impact of a single individual investigator working largely on his or her own within the language community, and imbued with the philosophy of the Himalayan Languages Project, is greater than any group effort. I do not want to say anything about the specific socio-political context of the Americas or Oceania, which is highly influenced by what is currently politically correct or fashionable in the Anglo-American world, but at least in the context of Asia and certainly in the greater Himalayan region this approach is best because meaningful cooperative community efforts can only begin after the foundation has been laid in the form of a holistic description. An individual investigator can blend into a language community more easily than a group, and he or she can only educate the community linguistically after he or she has gained enough mastery of the language to produce a holistic grammar, glossary and text corpus. The timing must be correct. After the grammar has been completed, the necessary foundation has been laid for community efforts at language preservation, and a relationship of trust has been built up and established.

Here it is relevant to point out one new cause for concern which has arisen from an unexpected quarter, i.e. the codes of ethics which are currently being drafted by funding organisations and international agencies. It is no coincidence that the drafts which are being touted about in the halls of academe are specifically inspired by the North American and Australian situation. Both North America and Australia, and to some extent Latin America, are the portions of the planet where European colonialism has been so successful that it has led to the virtual replacement of the resident peoples by intrusive populations. Especially in North America and Australia, where the native peoples have been marginalised and in some cases even driven to

extinction by newcomers, the issues of guilt and compensation have arisen, and relations have traditionally been sullied by distrust. At this stage, it is crucial to keep in mind that the societal and political issues which inspire ethical codes being drafted by agencies are peculiar to these parts of the world, where Anglo-American jurisprudence tends to transform complex moral issues into technical legal ones. An important lesson which we must learn from Zepeda's paper is that such codes of ethics do not make linguists more ethical; they merely impede research for the benefit of the community and the advancement of science. In her case, artificial demands impeded the work of a native American working in her very own native American community.

The situation in Eurasia and Africa is very different from that in North America or Australia. In fact, the specific situation in any given Asian country is utterly different from the next Asian country. Projects submitted for funding in Africa and Eurasia should not be subject to the same highly specific socio-political criteria which happen to be currently fashionable in North America and Australia, and to a lesser extent in Latin America. It spoils a naturally good relationship to require informants to sign working contracts with linguists. Descriptive linguists are not lumber companies, mining consortiums or colonial agents. The beneficial work of linguists should not be hampered by codes of ethics to which their research proposals must conform. We live in a topsy-turvy world when we spend our time dreaming up hurdles for linguists who strive to document and preserve endangered languages and cultures, whilst at the same time the threat to endangered language communities from lumber companies, mining consortiums and development schemes has not been effectively addressed. Descriptive linguists are on the same side as the language community, and they should not have to waste their time proving it. With all of the undocumented languages threatened with imminent extinction, linguists have better ways to spend their time.

Twenty years ago, the renowned Indo-Europeanist and Amerindian scholar Werner Winter told me to consider myself lucky to be working in the Himalayas, where the relationship between linguist and language community had not been spoiled as it had in North America. In Asia there were no clocks keeping time at sessions with informants. Now the situation in North America has only become more complicated. Yet well-intentioned policy makers are unwittingly making efforts to export the unhealthy American-Australian situation to the rest of the world. As the output of the Himalayan Languages Project has shown, linguistic work conducted in Asia tends to directly benefit the language communities concerned. It is inherent to the situation in most Asian countries that linguists naturally work in concert with the language communities and their representatives. The wholesome Asian situation can

only be spoiled by superfluous and culturally inappropriate regulation. The Himalayan Languages Project has been highly successful and has optimally benefitted the language communities for which it has worked. On the basis of my experience, it is my conviction that the success and the benefit of the research programme would only have been compromised if the project had been subjected to any of the various 'ethical' protocols that I have seen brandished about by various funding organisations and international agencies today.

(2) Holistic language documentation such as that produced by the Himalayan Languages Project provides the most complete and reliable documentation of an endangered language, both for the scholarly community and for posterity. Holistic grammars with text corpus and glossary provide the best and most reliable account of the language. Language is a complete organism in which regularities and linguistic phenomena at all levels of description are interwoven into one organic whole. As Antoine Meillet once observed, language is *un système où tout se tient*, i.e. a system within which everything has to do with everything else. Therefore, complete grammars, including a morphologically analysed natural text corpus and a bilingual dictionary, constitute the form of language documentation with the greatest value and utility. Whole descriptions tend to be more reliable than minor studies which attempt to document just the phonology or only syntax or discourse phenomena. Whole descriptions document the phonetics, phonology, morphophonology, morphology, morphosyntax, syntax and pragmatics of the language as well as a large portion of the lexicon, not to mention the lore, legends and oral traditions.

(3) Finally, a research programme producing holistic grammars is the approach that is of maximum benefit to the individual investigators who make up the research team. Each member of the team works on his or her own: One person to a language. Most members of the team are Ph.D. candidates, who traditionally start out in the Dutch system as A.B.D., i.e. 'All But Dissertation' in American academic parlance. The so-called course work traditionally takes the form of tailor-made individual guidance by the Ph.D. supervisor in the form of intimate sessions known as *privatissima*, granted on request, and intensive individual reading and study. A number of team members are post-doc scholars. In either case, the single criterion upon which the individual researcher will be judged is his or her original and substantive contribution to science in the form of the holistic document on the endangered language. The Himalayan Languages Project set-up, therefore, provides the best opportunity for an individual scholar to advance his or her scientific and scholarly career. The programme is to some extent self-managing and self-regulating because each member of the team has his or her own domain, i.e. a

language to himself or to herself. Therefore, he or she is motivated and allowed to be driven by his or her own academic and professional self-interest. This guarantees both the volume and quality of the output and reduces some of the overhead expense involved in monitoring the output.

There is one crucial strategic moment for the management, however. This is the moment of selecting candidates. We must face the fact that not everyone with formal academic training in linguistics is up to writing a holistic grammar of an endangered language. The daunting task of documenting everything possible about a hitherto undescribed language is decidedly not a job that every linguist is able to do. For one, a candidate must master or be able to quickly gain a sure-footed command of the contact language, whether this be Nepali, Mandarin Chinese, Hindi or Dzongkha. Holistic language documentation presents the greatest challenge and the intellectually most satisfying assignment for a young linguist. It demands both great versatility and the multi-faceted application of all of the investigator's analytical skills. The team members are recruited from anywhere in the world. The rigour of the selection process is even more crucial than the guidance which the project director can offer throughout the term of the individual researcher. The quality of the selection determines the raw material and the reservoir of talent within the research team and directly influences the outcome of each individual project. Even so, on the basis of the experience we have accrued, I would argue that the Dutch formula of four years for a Ph.D. candidate and three years for a post-doc scholar should be realistically changed into funding for a five-year and a four-year period respectively.

In conclusion, therefore, the determined and dedicated researchers of the Himalayan Languages Project deserve every form of continued support. Long-term research funding should be made available for initiatives which adopt this same approach to endangered language documentation. I argue here that this is the most cost-effective and multiply beneficial way to allocate research funding which has been earmarked for endangered languages.

Panel Discussion 1: Problem of Documentation

Participants: Marina Khasanova, Michael Krauss, Christian Lehmann, Toshihide Nakayama, Irina Nikolaeva, Tasaku Tsunoda, George van Driem
Co-chairpersons/editors: Fumiko Sasama (Osaka Gakuin University), Honoré Watanabe (Kagawa University)
Coordinator: Nobuko Yoneta (Osaka Jogakuin Junior College)

[The transcript of the tape recording of this panel discussion has been slightly edited for the sake of presentation in the present volume. Annotations, some of which were later provided by the participants, appear between brackets [...] as do information on genetic affiliation of languages and the country (countries) where they are spoken. Comments made in Japanese and translated into English are marked with (*) – editors/co-chairpersons.]

Watanabe:

First of all, I'd like to thank the seven panelists. We are honored to have you here in Kyoto. I thank you for the very interesting presentations we heard this morning. There were a wide variety of issues relating to documentation and description of languages, and especially of endangered languages.

I would like to mention that the issues may be categorized into two categories: that is, 'language internal' and 'language external'. Language internal issues refer to those in linguistics in a narrower sense; how to elicit sentences, how to analyze data, different presentations of language according to its type, and so forth. Language external issues include training of young students in field work that we heard from Professor Tsunoda, or budget, funding issues, or the ethical issues or social and/or political situation of the region that we are going into to do the research. Of course, it is very difficult to draw a clear line between the two. It is needless to say that the two are obviously inter-related, but maybe we can keep this in mind as we go on to the discussion.

Another aspect that we want to keep in mind is that we are going to be discussing about different issues of documentation and description of language, but we are especially concerned about endangered languages. We would like to try to elucidate what is different, what is special, and what is difficult about documenting and describing endangered languages, and what is different from working on so-called safe languages.