DOCUMENTING HIMALAYAN LANGUAGES: A CRITICAL REVIEW DEALING WITH TWO RECENT CONTRIBUTIONS

Mark Turin

Michael Noonan in collaboration with Ram Prasad Bhulanja, Jag Man Chhantyal and William Pagliuca. 1999. *Chantyal Dictionary and Texts* (Trends in Linguistics: Documentation, No. 17). Berlin: Mouton de Gruyter.

Werner Winter. 2003. *A Bantawa Dictionary* (Trends in Linguistics: Documentation, No. 20). Berlin: Mouton de Gruyter.

Mouton de Gruyter is one of a handful of European publishers who still specialise in linguistic research. A division of Walter de Gruyter, Mouton publishes conference proceedings, journals and monographs. Under review here are two recent volumes documenting endangered languages spoken in Nepal, *Chantyal Dictionary and Texts* by Michael Noonan et al. and *A Bantawa Dictionary* by Werner Winter. This review article is both a narrative about the nature of linguistic research and publishing in Nepal and a critical appraisal of these two new contributions to Himalayan linguistics.

Despite their superficial similarities in format, presentation and subject matter, the two books under discussion are quite different in scope. Winter's dictionary is of the Bantawa language spoken in Eastern Nepal, the most widely spoken of the Rai languages which make up an important subgroup of the Kiranti sub-division of Tibeto-Burman languages. According to the publicity material which accompanies the publication, "the dictionary, based on material obtained in the context of the *Linguistic Survey of Nepal*, concise though it is, stands out as the most voluminous of the few dictionaries and word lists for Rai languages that were hitherto published with English equivalents of native forms provided." Notwithstanding the wealth of data presented, however, the dictionary is of limited utility on account of the absence of supporting information for the lexical entries. Readers are left to puzzle out for themselves whether the lexical items are nouns, adjectives, verbs or adverbs; and are given no help in decoding Winter's chosen notation system or in understanding the difference between '=', '+' and a hyphen

'-' used to separate elements. We are also left to interpret subtleties such as the differences in meaning between the four Bantawa words translated as 'anywhere' (p. 53) and the three words meaning 'live charcoal' (p. 75). Nevertheless, the dictionary does contain almost 5,000 entries which will be of interest to scholars of Himalayan languages pursuing comparative lexical or phonological work.

Winter's short *Preface* raises two further interesting questions, one relating to linguistic politics and the other to the ethics of scholarly collaboration. The author states that the present volume was explicitly limited to 'native Bantawa vocabulary' and that the forms included are 'almost exclusively pure Bantawa' (p. v). While this goal is one shared by many descriptive linguists working in Nepal, I am intrigued by local reactions to this decision. Linguistic minority communities across Nepal are at present engaged in a battle for effective representation at a national level, and many language activists from these communities are busy preparing lexical collections of their own. In this process, as I have discovered first hand, the aim of the field linguist to record and preserve only the 'native' or 'pure' linguistic forms is often at odds with the desire of indigenous activists to bolster the size of their mother tongue corpus with loan words, usually from neighbouring languages or Nepali.[1] The aim of language activists, then, is more likely to be the compilation of a lexicon which appears impressive in the hope of greater recognition from political powers. This leads me to wonder how different the word list would have been, had a mother tongue Bantawa speaker been the author.

Furthermore, it is necessary to establish which factors are being used to determine linguistic nativism or purity. In the Thangmi language spoken in Dolakha and Sindhupalcok, for example, the indigenous word for 'shaman' is *guru* (Nepali *jhankri*). While younger and educated ethnic Thangmi who are literate in Nepali are aware that the term *guru* has an Indo-Aryan origin and that it is widely used in Nepali and Hindi in a range of meanings from 'teacher' to 'bus driver', older Thangmi shamans sincerely believe the term to be native to their Tibeto-Burman tongue. A linguistic fact, namely that *guru* is an Indo-Aryan loan word, may have no analytical value in the ethnolinguistic context of modern spoken Thangmi and may thus be rejected from an indigenous perspective. As a comparable example, one could think of telling an American social scientist that the word *anthropology* is not a native English term but

---

1 I have discussed this in my introduction to the *Nepali - Thami - English Dictionary*, published by Martin Chautari (Turin with Thami 2004).

rather a loan from Greek, and thus not to be included in any dictionary of 'pure' English.

The second issue relates to the inherently collaborative process of documenting a language which is not the mother tongue of the field linguist. Descriptive linguistic research, somewhat like traditional ethnography, requires the linguist to immerse him- or her-self in the socio-linguistic context to truly appreciate the nuances of the spoken form of the language under observation. Such exposure is only possible with considerable assistance from one or more native speakers of the language, typically individuals who are extremely patient and willing to spend hours explaining, repeating and correcting the linguist's child-like errors in the tongue he or she is attempting to decode. Of the various elements which make up linguistic research, the compilation of a lexicon lends itself most readily to being pursued in a genuinely collaborative manner. While studies of grammar, verbal morphology and phonology may appear very abstract to many mother tongue speakers of endangered languages who have no training in linguistics, the process of (and need for) collecting words is often self-evident. In fact, unlike grammatical analysis for which training in linguistics is a must and being a non-native speaker of the language may even be an advantage, in the collection and compilation of a word list or lexicon, a mother tongue language speaker – regrettably referred to as an 'informant' by most linguists – is at a distinct advantage. He or she invariably knows more words than the fieldworker does, and has a more complete vocabulary of the language at his or her disposal. The task of the linguist is then to compare and segment these lexical items which are produced from unelicited speech and controlled interview-like situations. Since lexicographical work is one aspect of linguistic fieldwork which is truly collaborative, linguists often choose to co-author dictionaries with their language teachers or 'informants.'

In the *Preface* to his *Bantawa Dictionary*, Professor Winter carefully acknowledges the assistance of two mother tongue Bantawa speakers, Mr. Tika Ram Rai and Dr. Novel Kishore Rai, the latter a distinguished linguist in his own right. While Winter notes that Dr. Rai 'went through our lexical corpus of Bantawa, to which he himself had substantially contributed by the lists included in his unpublished Pune dissertation' (p. v), *A Bantawa Dictionary* has nevertheless been published as a single-authored work.

Chantyal Dictionary and Texts addresses the issue of collaboration very differently, with the by-line reading 'by Michael Noonan, with Ram Prasad Bhulanja, Jag Man Chhantyal and William Pagliuca'. Moreover,

in the *Notes* with which the primary author concludes his introduction, Noonan writes:

> I would like to thank Ram Bhulanja for his untiring assistance in the production of the materials published here. All of the material was discussed with and/or checked by him: any remaining mistakes are entirely my responsibility. Jagman Chhantyal transcribed a number of the texts from tape. Bill Pagliuca was active in the research project in its early stages (p. 6).

By including two mother tongue Chantyal speakers as secondary authors of this important corpus of Chantyal textual and lexicographical material, Noonan acknowledges the impossibility of conducting such a project without the unflagging assistance of native speakers. In addition, joint authorship acknowledges that it is not only academic linguists for whom heavy and impressively-bound hardback texts are a source of prestige and cultural capital. After collaboration with a foreign linguist, mother tongue speakers of endangered languages may want, or even need, to show some tangible 'result' or authoritative 'text' to their own linguistic communities. All the better then, if such a book acknowledges a native speaker as being the co-creator of the work.

Chantyal is a Tibeto-Burman language spoken by an ethnic group of the same name who inhabit the Dhaulagiri zone of western Nepal. While the Chantyal population numbers around 10,000, no more than 2,000 still speak the Chantyal language (p. 1). The English-Chantyal dictionary, by far the most substantial component of this volume occupying 480 pages, is masterfully compiled. Following a brief explanation of the format of the entries (p. 7), the lexicon is arranged following English alphabetical order including transliterated Nepali forms for Chantyal words along with examples of many Chantyal forms in their spoken context. Some charming examples of recorded conversations include the Chantyal for 'Oh, Mickey [Noonan], move over a little!' (p. 9) and 'This Whiteman regards you as a son' (p. 331). The dictionary is replete with ethnographic asides, local colour and scientific names for local flora and fauna together with their uses, making this English-Chantyal dictionary a thoroughly useful resource for comparative research. The addition of six pages of kinship charts, a complete index of Chantyal morpheme and verb stems, and five transcribed, glossed and translated texts, referred to as 'discourses' by Noonan (p. 533), flesh out the already impressive lexicon with welcome ethnolinguistic data.

Both books under review were published in the same sub-series, *Trends in Linguistics*, in which the presentation of raw data takes precedence over contributions to general linguistic theory. Three of the twenty monographs in this sub-series relate to languages spoken in Nepal (the third being Ramawatar Yadav's excellent *A Reference Grammar of Maithili*, 1996), a statistic which warrants a moment's consideration. This fortunate, if apparently disproportional, focus on Nepal's linguistic heritage reflects three factors: the typological importance of and interest in Nepal's minority languages; a growing corpus of empirical linguistic data collected from within Nepal's borders; and the regional interests of the editors of the series. All monographs in the sub-series are further united by a minimalist, traditional typography, high quality paper and binding, and most importantly, an extremely high price. As a case in point, Werner Winter's *Bantawa Dictionary* retails at US$ 142 (over Nrs. 10,000), which for its 260 pages averages at over $0.5 per page. In comparison, Noonan's *Chantyal Dictionary and Texts,* which clocks in at 616 pages, is a steal at US$ 178.

The point here is not to lambast Mouton de Gruyter for their retail prices. Nor do I seek to criticise the authors for submitting their works to this highly-regarded publisher. Rather, this review provides an opportunity for assessing the publication and distribution of primary linguistic data. Unlike more discursive studies, in which the careful crafting of narrative is all important, both of the books under discussion are data-heavy and not primarily designed for 'reading'. In both cases, each printed page of the dictionary resembles a database print-out. Since both volumes are invaluable repositories of unique linguistic field data on little-known and endangered languages of Nepal, the publication of such information in expensive European hardback print editions may not be the most successful way of disseminating the findings.

First, the prohibitive cost of purchasing such a book effectively rules out all but the best-heeled Western university libraries, and by definition almost all scholars based in South Asia. Second, the presentation of lexical data on poorly-documented languages in a tabulated print edition limits the ways in which interested readers may interrogate or interact with the data. For example, I have spent a good few hours combing these two volumes for lexical information which may be of comparative interest for my own research, and have in the process compiled word lists of different classes. Had the data been presented in an online database format, employing Unicode diacritics for the encoding of IPA (International Phonetic Alphabet) glyphs, and supplemented with a simple

search and retrieval interface, the possibilities for linguistic examination and research would have been exponentially greater. Building such an interface is no longer an expensive or laborious process, as I have recently discovered through the compilation of a concise online word list.[2]

While I remain concerned about the lasting viability of the Internet as a data archive, and am aware of the need for refereed publications from established academic print houses which bring scholarly credit, I find it regrettable that the distribution of seminal work such as that produced by Winter and Noonan *et al.* should be so restricted in all of the developing world due to the high cost of specialised academic publications. Perhaps a two-pronged approach, combining a European or South Asian print edition with an online database would address both the need for scholarly recognition and wider dissemination. Given that the collection, cataloguing and storage of multimedia data is an increasingly indispensable component of the scholarly endeavour, it would be unfortunate if the medium of academic output did not embrace the new possibilities afforded by information technology. The recent establishment of *Himalayan Linguistics*, a refereed web journal and archive of grammars, dictionaries, and text collections specializing in the languages of the Himalayan region, edited by Michael Noonan, Carol Genetti and Tej Ratna Kansakar, is a positive move in this direction.[3] As for the concerns which foreign scholars may have about the standing and viability of publishing houses in Nepal, this too is changing. There are now more first-rate publishing possibilities than ever before, such as *Martin Chautari Books* and *Himal Books,* alongside the old favourites *Mandala Book Point* and *Ratna Pustak Bhandar*. These four publishers provide editorial assistance, proof reading and varying levels of peer review. The spread of the Web and the advent of professional imprints in Nepal are therefore changing the publishing landscape for Himalayan studies.

In this review, I have set out to question the practice of publishing research on Nepal's endangered linguistic heritage solely in the West and also to unearth some buried assumptions about the nature of collaborative work in field linguistics. While grammatical descriptions written in terse linguistic prose may be of little immediate utility to the communities whose languages they describe, word lists, annotated lexical materials and dictionaries are of very real benefit to endangered language communities

---

2 <http://iris.lib.virginia.edu/tibet/reference/dictionary/thangmi/index.php>

3 <http://www.uwm.edu/Dept/CIE/HimalayanLinguistics/>

at practical and symbolic levels. Only by empowering indigenous people and the mother tongue speakers of these endangered languages with the analytical tools needed for linguistic research, and by working collaboratively with such communities to ensure that research findings are locally accessible, can important languages such as Bantawa and Chantyal be rigorously documented and truly championed.

References
Turin, Mark with Bir Bahadur Thami. 2004. *Nepali - Thami - English Dictionary*. Kathmandu: Martin Chautari.

Yadav, Ramawatar. 1996. *A Reference Grammar of Maithili* (Trends in Linguistics: Documentation, No. 11). Berlin: Mouton de Gruyter.