Documenting One Language in a Multi-Lingual Community

CHRISTINA M. WILLIS
University of Texas at Austin

0. Introduction
Increasingly linguists agree that mono-lingual communities are not the norm (Woolard 1999). As such, the fact that most areas of India are multi-lingual is hardly surprising. The reported number of living languages in India varies from 114 languages, (Abbi 2001) to 415 (Gordon 2005). Including Hindi and English, there are twenty-three official languages recognized by the constitution of India. The official languages are spoken throughout the subcontinent alongside the many other un-official languages (and their dialects).

This paper examines the linguistic situation of the area in and around Dharchula, Uttarakhand1 a border town in the Indian Himalayas, which serves as the field site for my language documentation and description project2 of Darma, a Tibeto-Burman language (called Darmiya by Grierson 1967-68, and others). I will demonstrate the importance of employing a methodology for data collection that results in natural discourse, and how this discourse can highlight the effects of language contact and the frequency of borrowing, and codeswitching. While these language practices can make it challenging for linguists to tease apart indigenous linguistic forms from borrowed forms or segments of codeswitching, relying on natural discourse for analysis provides data that realistically reflects the practices of a multi-lingual community. I will also demonstrate that by employing multiple methodologies, a researcher can gain a more thorough understanding of the language. Additionally, working with the same community for an extended period of time brings a deeper understanding of the language and the linguistic practices of its speakers.

---

1 Formerly part of Uttar Pradesh, Uttarakhand was known as Uttaranchal until early 2007.
2 My research has been financially supported by The University of Texas at Austin (LAGR Award); Fulbright (IIE Scholarship & DDRA Fellowship); and National Science Foundation (Dissertation Improvement Grant: BCS 0236475, Anthony C. Woodbury, supervising PI). The Darma people and the people of the Dharchula community also supported my research by agreeing to allow me into their homes and lives to listen, learn, record, and ask questions. Without their cooperation this project would not have been possible; I cannot thank them enough.
The Community of Dharchula, India

The Darma people are one of three groups that comprise the Rang tribe (called Bhotiya by the Government of India). The Rang people live in the state of Uttarakhand, India in and around Dharchula. The town sits in a narrow valley above the Kali River, which serves as a natural border between India and Nepal. Dharchula is one of the last major communities before the Indo-Tibetan border to the north. I made multiple trips to Dharchula between 2002 and 2005; the town served as my base of operation during fieldwork.

The Darma traditionally migrate seasonally spending the winter months in and around Dharchula (elevation ~3,000 ft), and the summer months in the Darma Valley (elevation ~10,000-14,000 ft). The Darma Valley lies northeast of Dharchula, and comprises fourteen villages, which straddle the banks of the Dhauli Ganga (called the Darma River by some Rang people).

Before 1962, the Rang conducted trade at markets in Tibet; this practice ended abruptly in 1962 when a conflict between China and India closed the border, which forced many Rang to adopt a sedentary lifestyle. Those who took steady jobs ceased to participate in the annual migration. While the border reopened in 1992, few Rang have returned to the trade practice. During my research I found that while the majority of Rang people remain in the Dharchula area, few families could afford time off work to make the seasonal migration; those families who do often live at a subsistence level. Some people who grew up in the post-trade era now work as civil servants; they live away from the Dharchula community to fulfill their duties. This has resulted in a growing Rang diaspora across India. It is difficult to ascertain the exact number of Rang people, and the population of Darma people is merely an estimate. Based on the 2001 Census (Office of the Registrar General 2001), the Darma population is about 2,615.

In the Dharchula community there are a number of languages used in daily interactions, and most people are multi-lingual. The languages spoken belong to two genetic families: Indo-European (AKA Indo-Aryan) and Tibeto-Burman. For a researcher from outside the community, however, identifying all of the languages is problematic. This is due in large part to the fact that the people themselves disagree on what to call some of the languages they speak. In the following subsections, I will provide an overview of the languages identified by members of the community and discuss the status of each within the linguistic community.

1.1. The Tibeto-Burman (TB) Languages

There are three TB languages spoken by the Rang people in the Dharchula community Byansi, Chaudangsi, and Darma (some older Rang people also speak Tibetan). Byansi is reported to have three distinct varieties (S.R. Sharma 2001a, Trivedi 1991), but these have not been documented separately for comparison. Of the three languages, only Byansi has been described in a full grammar (Trivedi 1991). There are short sketches of Byansi (S.R. Sharma 2001a), Chaudangsi (Krishan 2001), and Darma (Krishan 2001); none of these sketches is deemed
complete by its author. While the Rang consider their individual languages distinct, they commonly refer to all three varieties as Rang boli or Rang lo, both of which literally mean ‘Rang dialect’ (boli is Hindi and lo is Darma).

While the classification of these three TB languages is not conclusive (Saxena 1997, S.R. Sharma 2001b), they are considered to be closely related (Hale 1982, Ruhlen 1991, Voegelin and Voegelin 1977). Ruhlen classifies the Rang languages as sisters to Rangkas under the Almora branch of the West Himalayan, Tibeto-Burman languages (1991:331). In other systems, Darma, Byansi and Chaudangsi, while not classified as sisters, are closely related (cf Hale 1982, Voegelin and Voegelin 1977).

1.2. The Indo-Aryan (IA) Languages
Like the rest of India, Uttarakhand is linguistically diverse. In this mountainous region, speakers of IA languages and speakers of TB languages often live in the same communities. While Hindi is used for some daily discourse and is the medium of instruction in many local schools, some people in the Dharchula area do not speak Hindi well. Those who do not speak Hindi rely on alternate IA languages as their lingua franca. In addition to the TB languages, there are at least three IA languages spoken within the Dharchula community and possibly four. When I asked which languages (other than Hindi) are used locally, they were identified as Pahari (pahar means ‘mountain’ in Hindi), Kumauni, and Nepali. In his description of Byansi, Trivedi reports three IA languages: Kumauni, Nepali, and Hindi (1991:1).

Some locals claim that the Pahari, Kumauni, and Nepali spoken in and around Dharchula are three distinct varieties, while others claim these are different names for the same language. Speakers of Nepali from Kathmandu say that the Nepali spoken in Dharchula is very different from the variety in the capital. Likewise, speakers of Kumauni from Pithoragarh say that the Kumauni in Dharchula is a different dialect. This latter claim is supported by the description of Kumauni in the online version of the Ethnologue (Gordon 2005), and by D.D. Sharma’s historical account of the languages of the region (1983). Some locals explain the situation by stating that neither Nepali nor Kumauni is spoken in Dharchula; rather the language is Pahari. This confusion about the linguistic scenario is compounded by the fact that the IA languages of Dharchula have not been the focus of linguistic study and have not been documented separately. For the purposes of this paper, however, I will follow Trivedi and define the local IA languages as Hindi, Nepali, and Kumauni.

---

3 The TB languages are generally not spoken by non-Rang people. The exception to this is the lower caste people who work as the servants of Rang families. While these lower caste people understand and speak the Rang languages, the Rang usually do not recognize them as part of the speech community.
4 This autonym ‘Pahari’ used by people in Dharchula for their language should not be confused with the different varieties of Pahari spoken in the state of Himachal Pradesh.
1.3. Language Status

In this region, the language of standard education is Hindi, which is also the official language of the local government. As a result, Hindi has the highest prestige within the Dharchula community as a whole. Within the Rang community the Byans people hold the most prestige and have the most economic resources (S.R. Sharma 2001a, Trivedi 1991). The Chaudangs is the smallest Rang tribe and is often considered a subset of the Byans. The assumption that the Chaudangs speak a dialect of Byansi (D.D. Sharma 1989) has recently been questioned due to the fact that Chaudangsi has not been described separately. Krishan (2001) provided a brief sketch of Chaudangsi, with the proviso that further work needs to be done to establish the relationship between Byansi and Chaudangsi. While the local perception is that Darma is the ‘original’ Rang language, the Darma people have the lowest socio-economic status within the Rang community. Unlike the Byans and Chaudangs who reside in Dharchula, the economic center of the area, the Darma spend their winters in smaller towns that do not have many opportunities for employment. As a result many Darma people survive on subsistence agriculture combined with meager earnings from selling traditional hand-woven rugs and blankets.

The social hierarchy of the Rang community manifests itself in a number of ways, the most salient to me as a researcher pertains to the language choices Rang people make in public. While visiting with Chaudangs and Darma shopkeepers in the local bazaar, I noticed, and it was reported to me, that Byansi speakers would frequently initiate a conversation in Byansi. Sometimes the non-Byans Rang person would respond in Hindi or Kumauni. I never witnessed a Darma speaker initiating a conversation with a Byans or Chaudangs person in Darma. I was told that is was because Byans and Chaudangs people do not understand Darma. Conversely, Byans people reported that Byansi is the lingua franca of the Rang tribes. This, however, is not the case. I encountered many Rang people who neither spoke Byansi nor considered it a lingua franca of the Rang people.

When I trekked to the villages in the Darma Valley I met a Byans woman, whom I call ‘Auntie’, who is married to a Darma man; they are part of the Rang diaspora. Many of her female in-laws were born and raised in Darma-speaking families, are married to Darma men who live in the Rang community, and speak Darma on a daily basis. Auntie, however, uses Hindi in her daily life, and does not speak Darma fluently. Despite this, Auntie participated in my research enthusiastically by initiating conversations and encouraging her Darma relatives to speak so that I could record them. Her in-laws, aware that my project was to document Darma specifically, chided Auntie for using other languages while I was recording. Even when prompted to speak in Darma, Auntie spoke in Byansi, English, and Hindi.

There are also English-medium schools in the Dharchula area. Increasingly locals view English as the language required to achieve success in modern society. While English has very high prestige and I sometimes used it during my fieldwork, few families use the language in the home.
Auntie’s language practice demonstrates that while people are adamant that there are three distinct varieties of Rang languages, speakers do not always recognize the differences. In general I found that Rang people viewed their languages as dialects in the pejorative sense of the term. Meaning they view Hindi as a language, but Darma as a ‘mere dialect’. Some speakers laughed at my project; they did not see how I could produce a descriptive grammar of a dialect because it is devoid of grammar. This perception of Darma as a dialect along with the higher status of the other Rang tribes led to situations where a group of Darma-speaking people would cede the floor to a Byansi-speaking person while I was recording. Because I do not speak Darma fluently, I was sometimes not aware of the situation until I sat down with my consultant to transcribe and translate the recordings. Fortunately, my primary consultant would alert me to places where someone had switched to Byansi.

2. Current Approaches to Language Documentation and Description

Currently, linguists rely primarily on a ‘discourse-centered approach’ to documentation and description (Sherzer 1987) rather than on direct elicitation where utterances are translated from the contact language into the target language. By focusing on naturally-occurring speech, the linguist can find and analyze words and structures that might not surface when sentences from one language are translated into another. Additionally, the texts gathered through this approach are culturally significant and provide context-based data. This methodology has become widely accepted, and it is the norm for linguists to base their descriptions of a language on natural discourse. It is also the norm for modern descriptive grammars to contain complete interlinearized texts so that ‘raw’ data is available to everyone (cf Epps 2005, LaPolla 2003, and others).

The languages of the Himalayas are found to share features with nearby IA languages. Some of these similarities can be attributed to the fact that the patterns found in the languages of South Asia form a distinct linguistic area (Masica 2005). One areal feature of South Asian languages is the converb construction. The converb is a non-finite particle that is found with verb stems and is used to concatenate multiple clauses under a single matrix verb. This type of discourse structure is found in both IA and Himalayan TB languages (e.g. Dolakha Newari as described by Genetti 1994).

While the presence of converb constructions is a feature of South Asian languages, the pattern found in Darma is slightly different than what is found in IA languages such as Hindi. In Darma, the converb construction is found throughout natural discourse (especially in historical narratives); it is used to advance the narrative and tie events together. As we see in example (1) below,  

---

6 Examples are written in a practical orthography where e = [ɛ]; ee = [e]; t = voiceless dental stop, t’ = voiceless alveolar stop, c = voiceless palatal stop, j = voiced palatal stop, 7 = glottal stop; [STOP]h = aspirated stop; ng = [ŋ]; x = voiceless palatal fricative; xh = voiceless uvular fricative; rC = retroflex consonant; r’ = alveolar tap; V = nasalized vowel.
several clauses can be concatenated under a single matrix verb (each converb\(^7\) is presented in bold and the matrix verb is underlined).

(1) [to \textit{7esa hee jo } \[7u su jo nini \]
[then like.this.H be.H that.H \[3SG ERG um \]
khurapat \[dimag \textit{ga-len } ju, \]
\[7u su gertho \]
mischief.\textit{idea.H mind.H do-CVB after} \[3SG ERG mill.\textit{flour} \]
mee \textit{po-len } ju, \]
\[7 i do \textit{tee e ki-ta-su} . \]
fire \textit{light-CVB after} \[\textit{police.E ERG um capture-CVB after} \]
\[j o nini, raja daro sar pu-kur’-su.] um \textit{king.H near deliver COMPL-take.away-PST} \]

‘Then it is like this, that he, um, after planning the mischievous idea, after he set the mill on fire, the police, um, after capturing (him), um, (they) took him away and delivered (him) to the king.’ (T0025: Kiti Phondar. 019)

Converb constructions like this are rarely found in elicited utterances. Those that do appear in elicited examples were usually produced after I provided a parallel construction in Darma. For example, in an elicitation session with my primary consultant I provided the sentence ‘She took the baby into the house and kept it there’ fully expecting a converb construction in the Darma version. Instead I got example (2) below. At the time I knew that converb constructions were commonly used, and after some coercion I was able to obtain example (3) below. My consultant then stated that the latter example was ‘better Darma.’

(2) \[7u su min xyeno song r’u \textit{xhe-su}, \]
\[3SG ERG small child house LOC house-PST] \]
\[7ido teer’ee ki-ta-su.] \]
then \textit{over.there COMPL-keep-PST} \]

‘She took the small child into the house, then (she) kept (it) there.’

(T0042: Elicited. 513)

\(^7\) Abbreviations used for the Darma data are as follows: person is indicated 1, 2, 3; ABL ablative; BEN benefactive; COMPL complements; COND conditional; CVB converb; DEM.NEUT demonstrative neutral distance; DEM.NONVIS demonstrative non-visible; E English word; EMPH emphatic; EMPRO emphatic pronoun; ERG ergative; FUT future; H Hindi word; INF infinitive; LOC locative; NOM nominalizer; NPT non-past; PL plural; POSS possessive; PST past; REL relative; SG singular.
Documenting One Language in a Multi-Lingual Comm.

(3) [[7u su min xyeno song r’u xhe-lan]
[[3SG ERG small child house LOC house-CVB]

teer’ee  ki-ta-su]
over.there  COMPL-keep-PST]

‘She took the small child into the house and kept it there.’ (T0042: Elicited. 514)

Without the natural discourse texts that I analyzed for the description of Darma, my grammar would not include an adequate description of the converb constructions as they are commonly used in narrative texts.

3. Challenges for the Researcher

Keating and Egbert (2004) discuss the nature of discourse in cultural terms. Different cultures have different norms for conversation. As an outside researcher, understanding what the norm is for the society being studied can be a challenge. Is the conversation following an appropriate pattern or are the speakers diverging from the pattern to assert power or show deference? In a multi-lingual speech community, understanding what the norms are can become even more complicated when the researcher is studying texts that have borrowed words, segments of codeswitching, interference from a prestige language, and so forth.

The fact that it is difficult for researchers to tease apart codeswitching, borrowing, interference, and contact phenomena is not new, but it is important that researchers acknowledge the limitations of their analysis (Woolard 2004:82). This is especially true when the researcher is neither from the community nor a native speaker of the contact language. In my own research, some of the understanding that I now have about conversational norms and borrowing and codeswitching are the product of working over the course of several years within the same community. I hope that as I continue to work with the Darma, my understanding of the language use and practice within the community will deepen.

One problem I faced in the Darma-speaking community was getting people to allow me to record. Often when people resisted being recorded, they would explain their reluctance by saying that I should record a ‘good Darma speaker’ and that I should not waste precious disc space on their conversation. The notion of a ‘good Darma speaker’ perplexed me for quite some time, and my consultants were unable to explain to me what people meant by this. Eventually, I was brought to a man deemed a ‘good Darma speaker’ who, in front of a large gathering of Darma, told a story and sang a song. I was surprised that the man was deemed a good Darma speaker because throughout his narrative he used lots of Hindi words and sizeable portions of the story were told in Hindi. The song turned out to have a lot of Kumauni in it.

So why was he deemed a ‘good Darma speaker’? Clearly my idea of a good speaker and the perception of the Darma community did not match up. To me a good speaker is someone whose stories contain Darma words and structures. It
wasn’t until later that I began to understand what it means to be a ‘good speaker’ in the Rang community. The act of story telling or leading songs requires innovation, wit, and great skill with words. The words do not need to be Darma, they just need to be woven together well. After I recorded the man who used lots of Hindi, another man took the floor. While his narrative contains less Hindi, his style was not as well received. In the text, he recapitulates the end of the story the previous man had told, and uses it to segue into the story of another famous Darma person. Just after he segues, a woman in the audience accuses him of conflating the two stories, which was perceived as evidence of his inferior style.

Factors such as language contact, language shift, language change, and codeswitching can result in usage that some speakers would not consider part of their language. Despite the use of borrowing and codeswitching by the skilled Darma speaker, when questioned directly about some borrowed structures, speakers would not recognize them as indigenous to Darma. This was especially true if the borrowed structure was juxtaposed with a Darma construction. Finding structures that are potential borrowings and comparing them to indigenous forms is not always easy. This challenge can be compounded by the fact that the borrowings may come from the contact language (i.e., the language used in interview sessions).

One example from my research involves a type of relative construction that I found in natural discourse called a correlative. Like the converb, the correlative is an areal feature of South Asian languages. I had reason to believe that the correlative was available in Darma based on D.D. Sharma’s comparative grammar of the Tibeto-Himalayan languages (1994:278). Structurally, correlative constructions are different from relative constructions. An utterance with a relative clause such as ‘the boy who is standing is tall’ is rendered ‘which boy is standing, that boy is tall’ in a correlative construction. The Darma example provided by D.D. Sharma is shown in (4) below.⁸

(4) khämi ra-yän idu de-yän.
  who come he go
  ‘Who(soever) comes he will go.’ (D.D. Sharma 1994:278)

During my research I found that speakers used the correlative construction in natural discourse. Examples are shown in (5)-(7) below.

---

⁸ This example is presented using the orthography employed by Sharma. The ä is schwa.
Confirming the correlative construction during elicitation sessions was not a problem. When I provided my primary consultant with a correlative construction in Hindi and asked for the Darma equivalent, I would get the exact same structure with Darma words. At one point I began to notice a construction that was not a
correlative, but appeared to function as a relative. Examples from natural
discourse are shown in (8)-(10) below.

(8) …kidang dee-nu mi...
… Kidang go-NOM person
‘…the people going to Kidang…’ (T0025: Kiti Phondar. 061)

(9) Dharchula xyung-nu wala jen ni-yang.
Dharchula sit-NOM one.H PL be-FUT
‘(They) will be people who sit in Dharchula.’ (= ‘They must be residents of Dharchula.) (T0032: Conversation. 125)

(10) hā wan-lan jo nini khami tuktu wan-je nongdi then arrive-CVB um who first arrive-COND later
wan-nu mi jen jati ga-ta.
arrive-NOM person PL food make-3.NPT
‘Then, arriving, um, whoever arrives first makes food for the people who
arrive later.’ (T0033: Alam Ceremony. 002)

In another elicitation session, I tried to obtain a relative clause from a
consultant who also speaks English. I thought that by not providing the Hindi
correlative, my chances of eliciting a relative clause would increase. When I
supplied the English prompt ‘the boy who is standing is tall’, however, my
consultant provided the correlative construction found in example (11) below. I
tried a different sentence and got the example in (12) below. I then provided my
consultant with an example of the structure I was looking for, and he provided
example (13) below.

(11) hadu jo ki-ne-n(u) ni-ni 7idu xyenu
DEM.NEUT REL.H COMPL-stand-NOM be-3.NPT DEM.NONVIS boy
bungnu ni-ni.
tall be-3.NPT
‘That one who is standing, that boy is tall.’ (T0042: Elicited. 373)

(12) t’eme bera ga-da 7idu filam su lee.
girl song do-3.NPT DEM.NONVIS Filam from be
‘The girl is singing. She is from Filam.’ (T0042: Elicited. 377)

(13) bera ga-nu t’eme filam su lee.
song do-NOM girl Filam from be
‘The girl who is singing is from Filam.’ (T0042: Elicited. 379)
When comparing the Hindi-like correlative in (12) to the relative structure in (13), my consultant declared that the relative construction was preferred. These examples show the importance of not taking every structure found in natural discourse at face value. This borrowed structure nearly obscured an indigenous structure that is still in use.

4. Conclusion
In this paper, by presenting the scenario from my own documentation project, I have shown the importance of the linguistic context of data obtained during the documentation and description of a language. I have also reinforced the assertion that a discourse-centered approach to documentation and description provides an efficient method to gathering contextual data, and that such an approach is especially valuable in a multi-lingual speech community. While the discourse-centered approach to data collection is the preferred method, I hope that I have also demonstrated that it is still imperative to further explore the structures obtained through natural discourse with direct elicitation. While the researcher should not rely solely on a translation approach, direct elicitation can be combined with a context-based methodology to gain a deeper understanding of speaker ideologies (e.g. which structures the native speaker deems ‘better’). Finally, the situation of my own fieldwork has brought to the forefront the fact that working with the same community for an extended period of time brings a deeper understanding of the language and the linguistic practices of its speakers.

References

Christina M. Willis
The University of Texas at Austin
Department of Linguistics
1 University Station, B5100
Austin, Texas 78712

cmwillis@mail.utexas.edu