elpublishing

Language Documentation and Description

ISSN 1740-6234

This article appears in: *Language Documentation and Description, vol 4*. Editor: Peter K. Austin

What elicitation misses: dominant languages, dominant semantics

DAVID BRADLEY

Cite this article: David Bradley (2007). What elicitation misses: dominant languages, dominant semantics. In Peter K. Austin (ed.) *Language Documentation and Description, vol 4*. London: SOAS. pp. 136-144

Link to this article: http://www.elpublishing.org/PID/053

This electronic version first published: July 2014

What elicitation misses: dominant languages, dominant semantics

David Bradley

1. Introduction[1]

When work in a language is conducted through the filter of a language of elicitation, many semantic fields may be reshaped due to the categories assumed by the fieldworker. This problem is particularly acute where the language being elicited is endangered and in the process of being replaced by a dominant language used as language of elicitation, and language consultants also speak the dominant language. Examples will be drawn from several semantic fields in some endangered Tibeto-Burman (TB) languages which contain more semantic categories than the dominant languages (Chinese, Thai, Nepali, etc.) which are replacing them. The examples cited include deixis, time ordinals, kin group classifiers, and special classifiers.

Another filtering issue which arises in translation is whether to translate literally or to make culturally appropriate adjustments to translations. Often such judgements result in choices which may convey a different meaning. This kind of problem is not restricted to translation from a dominant language or into an endangered language.

2. Deixis

Deictics may be divided into many more opposed categories, quite unlike those of the dominant languages, which have fewer distinctions. Gong[2] has a four-way system: proximal, medial, and two distals: to the south/west and to the north/east, as shown in Table 1. Gong is being replaced by Thai, which has a three-way system, proximal, medial and distal.

[1] I am very pleased to acknowledge the support of the Leverhulme Foundation in providing me with a Leverhulme Professorship and to ELAP for hosting me in this Professorship, and for the comments of colleagues at the translation workshop in February 2006 and subsequently. I also acknowledge the funding support of the Australian Research Council (A A59701122, A00001357, A59803475), UNESCO Intangible Cultural Heritage Section, and La Trobe University.

[2] Gong is a severely endangered TB language of the Burmic group spoken in two villages in western central Thailand.

David Bradley 2007. What elicitation misses: dominant languages, dominant semantics. In Peter K. Austin (ed.) *Language Documentation and Description, Vol 4*, 136-144. London: SOAS.

Table 1: *Gong and Thai deixis*

	Gong	Thai
'this'	heŋ^{21}	nii^{55}
'that (near)'	t^{h}ə^{21}	nan^{33}
'that (far, S/W)'	tuk^{55}	nɔɔn^{55}
'that (far, N/E)'	mɔk^{5}	nɔɔn^{55}

The Gong medial is not a 'by addressee' form, but refers to a distance closer than the two distals; the Gong distals may refer to things which are still within sight, and may include things which are closer than those which are referred to with Thai **nɔɔn**^{55}. Younger semispeakers of this severely endangered language use a system which lacks the two distal forms, and is thus even more restricted than the Thai system; in effect, a lowest common denominator system which avoids any distinction within distals and thus the semantic differences between the meanings of the distals in the two languages.

The Lisu[3] deictic system differs between dialects (Bradley 2003) but has a minimal system of eight deictics, as seen in Table 2. This can be compared with the deictic system of extremely closely related Lipo, which is much more strongly influenced by Chinese and has adjusted its deictic system to the two-term Chinese deictic system. Western Lipo has cognates of two Lisu deictic terms, proximal and the most basic distal, and has generalised the latter to cover all distal meanings, isomorphic with the Chinese system. Eastern Lipo has different lexical material, but the same system. Lipo has been in much longer and more intimate contact with Chinese over nearly a millennium; such isomorphism is a common long-term outcome of contact with a dominant language.

[3] Lisu is a TB language of the Central Ngwi group, spoken in China, Burma, Thailand and India.

Table 2: *Lisu, Lipo and Chinese deixis*

	Lisu	W Lipo	E Lipo	Chinese
'this'	$t^hø^{33}$	t^he^{44}	$ħe^{55}$	zhè
'that (by addressee)'	$ʔɑ^{55}\ t^hø^{33}$	go^{44}	$k^hɯ^{55}$	nà
'that (same level)'	go^{33}	go^{44}	$k^hɯ^{55}$	nà
'that (higher level)'	$nø^{33}$	go^{44}	$k^hɯ^{55}$	nà
'that (lower level)'	$dø^{33}$	go^{44}	$k^hɯ^{55}$	nà
'that (same level, far)'	ko^{55}	go^{44}	$k^hɯ^{55}$	nà
'that (higher, far)'	$nø^{55}$	go^{44}	$k^hɯ^{55}$	nà
'that (lower, far)'	$tɕø^{55}$	go^{44}	$k^hɯ^{55}$	nà

Despite the richness of the Lisu deictic system, which has even more terms in some dialects, descriptions of Lisu by Chinese linguists have missed many deictic terms. Even the best Chinese descriptions have only four basic terms, and lack the medial (by addressee) form as well as the three intensifier high tone distal forms. This highlights another problem: incomplete descriptions of systems which the outsider linguist does not expect to find and does not attempt to elicit.

3. Time ordinals

Time ordinals in a variety of TB languages have up to eight lexical preceding and following day and year ordinals (8, 7, 6, 5, 4, 3, 2 days ago, yesterday, today, tomorrow, 2, 3, 4, 5, 6, 7, 8 days hence, and likewise for years). This is far more than available in any dominant language of the area, and is completely missed in many descriptions. For example, Chepang[4] has a ±8-term system (Caughley 2000); many other TB languages have up to ±6-term systems (Kiranti, Michailovsky 2003; Ngwi groups, notably Lisu, Bradley 1994, 2006). The TB time ordinal words have widespread cognates across TB. Interestingly, the time ordinals have nothing to do with cardinal numerals. Those for the closest times ('today', 'this year', etc.) are the most lexicalised; for more distant times, greater regularity is observed. Fossilised cognate lexical material is often preserved within time ordinals.

[4] Chepang is an endangered TB language of the Mahakiranti group, spoken in central Nepal.

Table 3: *Chepang day and year ordinals*

Day din (Nepali)	year bar.sa (Nepali)/yat.ko? ryoh (archaic)		
8 days before	nik.nəm/nip.nəm	8 years before	yat.ko? ryoh
7 days before	sik.nəm/sip.nəm	7 years before	yat.ko? k^hyoh
6 days before	k^hik.nəm/gip.nəm/bik.nəm	6 years before	yat.ko? syoh
5 days before	?ik.nəm/ŋhi.nəm/rak.nəm	5 years before	yat.ko? tyoh
4 days before	k^huk.nəm/gak.nəm/pu.nəm	4 years before	yat.ko? syoh
3 days before	kyam.nəm/?ak.nəm	3 years before	yat.ko? ts^hyoh
2 days before	tsit.nəm/bar.nəm	2 years before	yat.ko? teh
yesterday	yoh	last year	teh
today	ten	this year	nek
tomorrow	syaŋh	next year	ɲam.pu?/ɲam.p^{h}u
2 days hence	tsit.səy/tis.səy	2 years hence	k^hak.pu
3 days hence	kyam.səy/lik.nəm	3 years hence	?ik.pu
4 days hence	k^huk/k^huk.səy/rik.nəm/pu.sə	4 years hence	k^hik.pu/bik.pu
5 days hence	?ik.səy/?uk.nəm/ŋhi.səy	5 years hence	lik.pu
6 days hence	k^hik.səy/bik.səy	6 years hence	gik.pu
7 days hence	sik.səy/tsik.səy	7 years hence	sik.pu
8 days hence	nik.səy	8 years hence	nik.pu

Many Chepang day ordinals preserve a cognate **nəm** of TB ***ni** 'day' which is replaced in nominal use by the Nepali loanword din but which survives in forms for 2 to 8 days before, and variably in forms for 3 to 5 days hence. The Chepang cognate of TB ***nik/niŋ** 'year' is preserved, but only in 'this year'. We may also note various other interesting phenomena, such as the reduced form of 'tomorrow' seen in forms for 2 to 8 days hence, and the use of part of the archaic 'year' form as a productive prefix in the years before forms.

Chepang is in the process of losing its TB numerals; these are being replaced by Nepali numerals. Chepang preserves just enough archaic numerals with TB cognates to show that these have nothing to do with the time ordinals: yat 'one', nis 'two', sum 'three', pləj 'four', po.ŋa 'five'. Thus what one might think of as more basic areas of lexicon (cardinal numerals, 'day' etc.) are replaced, while unusual lexical material is preserved in a much more elaborate time ordinal system than found in the dominant language.

Time ordinals are often omitted from descriptions of languages filtered through the prism of a local dominant language which has fewer such terms. Furthermore, some languages which are less endangered than Chepang have lost much of the lexical richness; compare closely-related Newari, a TB language under strong Indic influence for millennia, and without lexical time ordinals going beyond those also found in Nepali.

4. Kin group classifiers

Kin group classifiers are numeral classifiers which co-occur with a numeral and refer to specific groups of relatives within one family (such as a group which includes a mother and one or more of her children). These are found in most languages of the Ngwi (Loloish, Yi) Group of Tibeto-Burman (Bradley 2001), but are absent from all other languages of the area and from all other languages of the world. The exact semantic value of the kin group classifiers differs from language to language. Of course many languages in other parts of the world have other lexical devices for groups of kin, but not as part of the numeral classifier system.

One such system is Lisu, as exemplified in Bradley (1994, 2001, 2006).

Table 4: *Lisu kin group classifiers*

'siblings/cousins/spouses'	ʂl^{55}
'mother and children'	mɑ^{55} lɑʔ^{21}
'father and children'	pɑ^{55} lɑʔ^{21}
'grandparent and grandchildren'	pi^{55} liʔ^{21}
'great-grandparent and great-grandchildren'	mi^{55} liʔ^{21}

These forms are not directly derived from kin terms, and combine directly with a preceding cardinal numeral, as in the noun phrases shown in (1) - (3) below.

(1) sɑ^{33} mɑ^{55} lɑʔ^{21}

three mother.children
'three people (including a mother and two children, etc.)'

(2) ji^{55} wɑ^{21} li^{33} pɑ^{55}lɑʔ^{21}

They four father.children
'theyfour (a group including a father and three children, etc.)'

(3) wɑ^{33} nu^{21} ȵi^{21} ʂl^{55}

we INCL two sibling.cousin.spouse
'we two, you (my sibling/cousin/spouse) and I'

These forms are highly polysemous; for example, (2) must include a father and three other people from the same family, but the other three people could be three of his children, or two of his children and the spouse of one of the children, or his wife (the mother of the children) and two of their children, or the mother, one of their children, and the spouse of that child. From the

perspective of a person in a particular generation, two-generation groups could include oneself and one's father or oneself and one's children or one's grandfather and one's father or one's son and one's grandchildren, and so on. Those languages in the most intensive contact with unrelated dominant languages have lost this category; for example Lahu[5] lacks it, perhaps due to intensive contact with various Thai languages over a very long period.

Again, most descriptions of the Ngwi languages have completely missed the kin group classifier category; but they are easily found, once one knows what to look for. They have now been found in all branches of Ngwi and appear to be one of the shared characteristics of this subgroup of TB.

5. Special classifiers

Special classifiers are numeral classifiers which may occur instead of the general classifier[6] or a more specific classifier only with certain numbers. One is the Nosu[7] classifier for humans, which occurs only with numerals greater than two (or greater than one in some dialects). Some years ago I had the instructive experience of eliciting numeral classifiers for humans in various Nosu dialects, in the presence of a leading Chinese expert on Nosu, and I discovered that the Sondi dialect uses the general classifier **ma**^{33} for humans only with 'one', while the standard Shengza dialect uses it with 'one' and 'two', before shifting to the human classifier **jɔ**^{33} with 'two' and above in Sondi and with 'three' and above in Shengza. After 30 years of working on the language, he did not know about this dialect difference, which is of course absent from Chinese.

Instead of the normal general classifier (Lisu **mɑ**^{33}, Lahu **ma**^{21}), the Lisu and Lahu special general classifiers occur with numerals that have socio-religious significance, especially seven and nine, and in some dialects also three. Most things to do with offerings and other religious activities come in groups of seven or nine, and some in threes. The form is **lø**^{21} in Lisu and **lɛ**^{53} in Lahu; it is cognate with the general classifier in some other Ngwi languages such as Nisu, Nasu and Hani and with the classifier for round things in Burmese. These all derive from a reconstructed form ***lum**^{2}. Bible translations do not use these special general classifiers in Lisu and Lahu; whether this is

5 Lahu is a TB language of the Central Ngwi group, spoken in China, Burma, Thailand and Laos.

6 A general classifier is the unmarked numeral classifier, which occurs after a numeral with a wide variety of noun heads. A special general classifier is restricted to occurrence after certain numerals.

7 Nosu is a TB language of the Northern Ngwi group, spoken in south-western China.

an attempt at simplification, an avoidance of items with 'pagan' connotations, or simply an error is not clear.

The forms of the general, special general and other numeral classifiers differ considerably between Ngwi languages, because the entire numeral classifier system has only developed over the last millennium in this group of languages. The numeral classifier system in related Chinese is also an innovation, but a somewhat older one. This reflects the well-known tendency for areal similarities to develop and spread from a dominant language into minority languages.

6. Pitfalls of cultural translation

Eco (2004) advocates the replacement of culture-specific references in the language being translated from with an equivalent parallel reference in the language being translated into. Such references may be completely different in their meaning. Specific-purpose translations also often need to express referents or concepts from a dominant language which are absent from the language being translated into.

Translations with religious aims may deliberately avoid terms associated with traditional religion. One example is the Hebrew term **la'ana** 'wormwood (*Artemisia spp.*)' which is not translated with the correct Central Lisu term **lo^{35} k^{h}wɑ^{21}** in any version of the Lisu Bible, but instead variously translated either as 'bitter thing' or more often as a completely different plant, **ʂl̩^{35} ʂl̩^{21}** Chinese goldenthread (*Coptis chinensis*)'. Why? Wormwood is an essential plant very often used in a large number of non-Christian Lisu religious ceremonies and as a herbal medicine; Chinese goldenthread is a fairly similar-looking plant which also has medicinal uses, like the biblical wormwood, and also tastes very bitter. One other problem with the Lisu form for wormwood is that there are dialect differences which might make it difficult for all speakers to understand the referent: Northern Lisu **ji^{35} k^{h}wɑ^{21}** or **jo^{35} k^{h}wɑ^{21}**, Southern Lisu **li^{35} k^{h}wɑ^{21}** and so on. By contrast, the Lisu word for Chinese goldenthread is the same in all dialects.

Translationese also often requires the creation of new strategies for coining lexical material. For example, abstract nominals may be absent from a language, but a way of coining them is created and may become more productive through time. In Lisu, for example, abstract nominals are now created by an inversion strategy first devised by Bible translators in the 1920s. Lisu has very many body part plus verb compounds, such as **ni^{35} nø^{33}** literally 'heart want' meaning 'to love'. A couple of these were inverted to give new abstract nominal forms such as **nø^{33} ni^{35}** 'love' for Bible translation. While some speakers still reject most additional coinages, others gleefully create more, all with 'heart' in second position, even when there is no corresponding

compound with 'heart' preceding the verb: **dzɑ^{21} ni^{35}** 'eating', in the absence of a compound **ni^{35} dzɑ^{21}** 'heart-eat'.

Various kinds of translation-based strategies can thus have extensive effects on the structure of a dominated language and the understanding of pre-existing and newly-introduced concepts and their cultural associations.

7. Conclusion

Even relatively language-aware speakers of an endangered language may encounter difficulty in translating lexical material for which there are no categories in the dominant language. Longer-term presence in the speech community, use of text-based materials and progression toward use of the language itself in collecting data may overcome this problem. The aim of this paper is not to give a detailed description of the specific phenomena cited; these are discussed in more detail in the references cited.

In those languages which have already undergone intensive contact with a dominant language lacking a category, expected lexical forms may have been lost, as in the case of Lahu which lacks kin group classifiers when all surrounding related languages have them. Alternatively, fossilised forms may be preserved, as in Chepang, which keeps cognate material for time words such as 'day' and 'year' in time ordinal forms, but replaces them with loanwords in their core meanings.

It is also possible that mistranslations may be partly deliberate, as seen in section 6, or may reflect long-term failure by outsiders, including linguists, to observe forms and their meanings such as those discussed in sections 2 to 5. It is also quite likely for a fieldworker with limited experience in an area to miss the richness in many areas of language structure, such as those briefly outlined here. This is one of many reasons why parachute linguistics does not work well.

References

Bradley, David 1994. A dictionary of the northern dialect of Lisu (China and Southeast Asia). Canberra: Pacific Linguistics C-126.

Bradley, David 2001. Counting the family: family group classifiers in Yi Branch languages. *Anthropological Linguistics* 43/1: 1-17.

Bradley, David 2003. Deictic patterns in Lisu and Southeastern Tibeto-Burman. In David Bradley et al. (eds.) *Language variation: papers on variation and change in the Sinosphere and in the Indosphere in honour of James A. Matisoff*, 219-236. Canberra: Pacific Linguistics 555.

Bradley, David (ed.) 2005. Language endangerment in the Sinosphere. *International Journal of the Sociology of Language* 173.

Bradley, David, with Edward R. Hope, James Fish and Maya Bradley 2006. *Southern Lisu dictionary*. Sino-Tibetan Etymological Dictionary and Thesaurus Monograph No. 4. Berkeley: STEDT.

Bradley, David and Maya Bradley (eds.) 2002. *Language endangerment and language maintenance*. London: Routledge Curzon.

Caughley, Ross 2000. *Dictionary of Chepang, a Tibeto-Burman language of Nepal*. Canberra: Pacific Linguistics 502.

Eco, Umberto 2004. *Mouse or rat? translation as negotiation*. London: Weidenfeld and Nicholson.

Michailovsky, Boyd 2003. Suffix runs and counters in Kiranti time ordinals. In David Bradley et al. (eds.) *Language variation: papers on variation and change in the Sinosphere and in the Indosphere in honour of James A. Matisoff*, 237-251. Canberra: Pacific Linguistics 555.