

A Construction Morphology Analysis of the Tibetan Case System

Elya

1 Introduction

“Tibetan” is often used by laypeople to refer to a hypothetical, singular language of Tibet, in much the same way that the complex situation of the Sinitic languages is often boiled down in common parlance to just “Chinese.” However, like their distant Sinitic cousins, the Tibetic languages are in reality a complex host of dialects and languages of varying degrees of mutual intelligibility spoken across the Tibetan Autonomous Region, as well as neighboring regions of China including Sichuan, Gansu, Yunnan, and Qinghai and the neighboring countries of Nepal, Bhutan, Pakistan, and India. Collectively, these languages are often referred to as བོད་སྐད་ (bod-skad /pʰøːl.kɛːl/), with བོད་ (bod /pʰøːl/) being the endonym for Tibet and སྐད་ (skad /kɛːl/) being the word for “language,” generally used in conjunction with the name of a region to indicate the tongue spoken there, as in ཨ་མདོ་སྐད་ (a-mdo-skad /am.l.to.l.kɛːl/) “the language of Amdo,” ཁམས་སྐད་ (khams-skad /kʰam.l.kɛːl/) “the language of Kham,” or དབུས་གཙང་སྐད་ (dbus-gtsang-skad /yː.l.tsan.l.kɛːl/) “the language of Ü-Tsang” (which is also known as Central Tibet). This last language is often conflated with ལྷ་ས་སྐད་ (lha-sa-skad /la.l.sa.l.kɛːl/ or /ɛː.l.sa.l.kɛːl/) “the language of Lhasa,” and often dubbed “Standard Tibetan” or གཞུང་སྐད་ (gzhung-skad /ɕuŋ.l.kɛːl/).

I will henceforth refer to the major dialect groups as Lhasake, Amdoke, and Khamke. It is the first of these which this paper will focus on, and it is this name, or “Lhasa Tibetan,” by which I will refer to the language; however, it is worth noting that many speakers in the diaspora use a chimera of dialectal features when speaking “Standard Tibetan” which I may make note of from time to time.

This paper will outline a construction morphology analysis of the case system found in Lhasake, laying out novel descriptions of the various allomorphs of three of its six cases (one other being unmarked and the remaining two being monomorphic). These include the ergative, genitive, and the oblique.

Lhasake, like other Tibetic languages, is an “ergative” language, meaning that unlike most European languages it does not morphosyntactically distinguish between the nominative and accusative cases but instead between the ergative and the absolutive. The ergative is primarily used to mark the subject of transitive verbs but is also used here as a marker of the instrumental, indicating the tool or means by which an action is done. The oblique functions as a combination of the dative and locative cases, used similarly to the English prepositions “at,” “in,” or “on” as well as, at other times, “to” and “for.” Furthermore, it marks the indirect object of a verb and can be used in certain constructions to indicate the purpose of an action. Finally, the genitive generally subordinates one noun to another, indicating not just possession but also origin, material, etc.

2 Phonotactics

A brief explanation of the structure of the Tibetan syllable is required in order to explain certain phenomena we see in the case system. However, we will withhold from diving into an in-depth discussion of Lhasake phonology. Instead, it suffices to make three notes: Lhasake distinguishes between four tones in monosyllabic words but only two in polysyllabic words (and these only in the initial syllable). Furthermore, it distinguishes between short and long vowels, but this is only weakly distinctive.

Our third and most important note is this: while the classical language allowed for མབནདལའལལལ (/m b n d s l r ŋ g y/) in the coda—where current scholarship considers them to have had the realizations, [m p n t s l r ŋ k x]¹—Lhasake has reduced these to མབདལ (/m p ŋ k/) with the realizations [m p ŋ ʔ].² The lost codas have generally had the effect of changing the quality, length, or nasalization of the preceding vowel.

DeLancey 2003 lays out six syllable patterns in Lhasake: CV, CVV, CVN, CVVⁿ, CVC, and CVʔ. He describes these as the short open, long open, nasal-final, long nasalized, obstruent-final, and glottalized rhyme patterns respectively (with the last being occasionally realized as a long vowel with contour tone). For our purposes, it is important to note that historically (and orthographically) bimoraic syllables remain bimoraic in Lhasa Tibetan, and the same applies to monomoraic syllables.³

3 Ergative

Much of the work of documenting and analyzing Lhasa Tibetan has been done by the scholar, Scott DeLancey, who in “Ergativity and the Cognitive Model,” laid out the semantics of the Tibetan ergative. In Lhasake, the ergative is frequently used to distinguish between volitional and non-volitional subjects. The Tibetic languages broadly feature a lexical split between volitional and non-volitional verbs that appears, in many cases, to descend from a regular feature of verbal inflection in their common ancestor. Specifically, pairs like མོར (skor /ko:l/) and འམོར ('khor /kʰo:l/), both meaning “to turn,” differ in terms of volitionality (the first being volitional, the latter non-volitional) and seem to derive from a common root.⁴ However, whatever process gave rise to these pairs is no longer productive. Instead, the ergative case is used: subjects of transitive verbs always take the ergative, whereas intransitive subjects can take either the ergative or the absolutive depending on volitionality.

According to traditional understanding, the form of the ergative (as well as the other cases) is determined by the last letter of the preceding syllable. After དལ it takes the form གིས (gis); after དལས, the form ཀྱིས (kyis) (this form also appears sometimes after other letters when a historical ད was present; these were dropped during spelling reforms that took place over a thousand years ago); after རལས, the form གྱིས (gyis); and finally ས (s) in open syllables.

In Lhasa Tibetan, these simplify down to two forms: /ki/ གིས ཀྱིས and, for our purposes, the more complex form represented by ས. This latter ending generally has two effects, the first being relatively straightforward. If the preceding syllable ends in /a o u/,

1. Nathan W. Hill, “Latin h- and Tibetan 'a : a reply to Schuessler” ().

2. Scott DeLancey, “Lhasa Tibetan,” in *The Sino-Tibetan Languages* (Routledge), 270–288.

3. DeLancey.

4. Scott DeLancey, “Ergativity and the cognitive model of event structure in Lhasa Tibetan,” *Cognitive Linguistics* 1, no. 3 (1990): 289–322, <https://doi.org/doi:10.1515/cogl.1990.1.3.289>, <https://doi.org/10.1515/cogl.1990.1.3.289>.

then these are fronted to / ε y \emptyset /. The second effect is slightly more convoluted: historically, the coda ʌ was lost, yielding what DeLancey describes as a “glottalized rhyme” which can surface as either a long vowel with contour tone or as a glottal stop. For example, the word for “I” is ɕ (nga /ŋaː/), and its ergative form is ɕʌ (ngas /ŋɛːʌ/ /ŋɛːʌ/). For an example of the more common form, /ki/, we can look at a word like ʁiɕ (ming /miŋː/), meaning “name,” which has the ergative form, ʁiɕʌ (ming-gis /miŋːgiː/). We can analyze these two modern forms as essentially arising under two conditions: for words ending in a bimoraic syllable, the form /ki/ is used, while for words ending in a monomoraic syllable, the form / [+FRONT]ʔ/ is used.

Following the notation for schema and sub-schema outlined in Van der Spuy 2017, the general schema for the ergative is given as (1) with (2) and (3) representing the two sub-schemas.⁵

1.

$$\begin{aligned} &< /X/\omega_i \leftrightarrow [N_{\text{ABS}}]_i \leftrightarrow [ABS[SEM]]_i > \\ &\approx < /XY/\omega_j \leftrightarrow [N_{\text{ERG}}]_j \leftrightarrow [ERG[SEM]]_j > \end{aligned}$$

2.

$$\begin{aligned} &< /X\sigma_{[\mu,\mu]}/\omega_i \leftrightarrow [N_{\text{ABS}}]_i \leftrightarrow [ABS[SEM]]_i > \\ &\approx < /Xki/\omega_j \leftrightarrow [N_{\text{ERG}}]_j \leftrightarrow [ERG[SEM]]_j > \end{aligned}$$

3.

$$\begin{aligned} &< /X\sigma_{[\mu]}/\omega_i \leftrightarrow [N_{\text{ABS}}]_i \leftrightarrow [ABS[SEM]]_i > \\ &\approx < /X[+\text{FRONT}]P/\omega_j \leftrightarrow [N_{\text{ERG}}]_j \leftrightarrow [ERG[SEM]]_j >^6 \end{aligned}$$

As seen in the sub-schemas (2) and (3), this notation indicates that the form of the ergative relies on the number of morae in the final syllable. Conveniently, this exact same be applied (in essentially the exact same manner) to the genitive and the oblique.

4 Genitive

The genitive in Tibetan generally takes a form homophonous to the ergative. However, the form found in words ending in monomoraic syllables phonemically ends in a long vowel rather than the glottal stop. These are, as mentioned, usually both realized as long vowels and, owing to Lhasa Tibetan’s reducing non-initial tonemes, there is generally no phonetic distinction made.

Orthographically, the genitive is nearly identical to the ergative. After ɕ it takes the form ɕi (gi); after ɕʌ , the form ɕi (kyi) (again, this form also appears sometimes after other letters when a historical ɕ was present); after ɕʌ , the form ɕi (gyi); and finally ɕi (‘i) in open syllables.

Returning to the first-person pronoun, its genitive form is ɕi (nga-‘i /ŋɛːː/). Contrastingly, the genitive of the word for “name” is ʁiɕi (ming-gi /miŋːgiː/) which as you can see is identical to the word’s ergative form.

5. Andrew van der Spuy, “Construction Morphology and inflection,” *Lingua* 199 (2017): 60–71, ISSN: 0024-3841, <https://doi.org/https://doi.org/10.1016/j.lingua.2017.07.010>, <https://www.sciencedirect.com/science/article/pii/S0024384117300098>.

1.

$$\begin{aligned} &< /X/\omega_i \leftrightarrow [N_{\text{ABS}}]_i \leftrightarrow [ABS[SEM]]_i > \\ \approx &< /XY/\omega_j \leftrightarrow [N_{\text{GEN}}]_j \leftrightarrow [GEN[SEM]]_j > \end{aligned}$$

2.

$$\begin{aligned} &< /X\sigma_{[\mu, \mu]}/\omega_i \leftrightarrow [N_{\text{ABS}}]_i \leftrightarrow [ABS[SEM]]_i > \\ \approx &< /Xki/\omega_j \leftrightarrow [N_{\text{ERG}}]_j \leftrightarrow [GEN[SEM]]_j > \end{aligned}$$

3.

$$\begin{aligned} &< /X\sigma_{[\mu]}/\omega_i \leftrightarrow [N_{\text{ABS}}]_i \leftrightarrow [ABS[SEM]]_i > \\ \approx &< /X[+\text{FRONT}][+\text{LONG}]/\omega_j \leftrightarrow [N_{\text{GEN}}]_j \leftrightarrow [GEN[SEM]]_j > \end{aligned}$$

The genitive and ergative being largely homophonous, it is no surprise that their schemas are nearly identical. However, if we look to the oblique, it is striking how this system pervades the Lhasa Tibetan case system.

5 Oblique

Lhasa Tibetan, more so than many other dialects and languages in the Tibetic family, has truly lost most of its historical codas. This is evident in the two liquid codas— r (l)—which are generally both lost, yielding long vowels in their place (with r triggering vowel fronting just like r). This sound change is not fully complete, meaning that a coda r may be realized as /r/ or /:/.

The oblique orthographically and historically had two forms, la and r which, as you might well predict, appeared after closed and open syllables respectively. Owing to Lhasa's loss of coda liquids, the latter form is largely retained as the lengthening of the final vowel, much like the genitive, albeit without the vowel fronting associated with that case.

To return to our trusted examples, the oblique form of the first-person pronoun is ngar (ngar/ŋa:l/). In contrast, the oblique of the word for “name” is ming-la (ming-la/minj.la l/).

1.

$$\begin{aligned} &< /X/\omega_i \leftrightarrow [N_{\text{ABS}}]_i \leftrightarrow [ABS[SEM]]_i > \\ \approx &< /XY/\omega_j \leftrightarrow [N_{\text{OBL}}]_j \leftrightarrow [OBL[SEM]]_j > \end{aligned}$$

2.

$$\begin{aligned} &< /X\sigma_{[\mu, \mu]}/\omega_i \leftrightarrow [N_{\text{ABS}}]_i \leftrightarrow [ABS[SEM]]_i > \\ \approx &< /Xla/\omega_j \leftrightarrow [N_{\text{OBL}}]_j \leftrightarrow [OBL[SEM]]_j > \end{aligned}$$

3.

$$\begin{aligned} &< /X\sigma_{[\mu]}/\omega_i \leftrightarrow [N_{\text{ABS}}]_i \leftrightarrow [ABS[SEM]]_i > \\ \approx &< /X[+\text{LONG}]/\omega_j \leftrightarrow [N_{\text{OBL}}]_j \leftrightarrow [OBL[SEM]]_j > \end{aligned}$$

It is worth noting that in Lhasa Tibetan (and many other dialects), the oblique is often reduced to just ལ, with ར appearing in fossilized grammatical structures. The same, to a lesser degree, goes for the genitive which is often realized in its གི (gi) form regardless of the weight of the preceding syllable.

There exist a number of somewhat archaic forms of the oblique which may be found fossilized in various constructions throughout Tibetan grammar, but the most important for our purposes is the class of forms with the consonant /t/: ཏུ ཏུ ལ (tu du su /tu tu su/). Conveniently, this form is identical to the more common oblique when the preceding syllable is open, taking the form of ར (r). However, this does introduce a wrinkle to our nice, neat mora-based system. We cannot easily account for the fact that the historical coda ལ yields a different form of the affix despite phonemically having merged with the other coda coronal, ར. This would require us to introduce some lexical categories: certain words ending in “glottalized” rhymes would take either the ཏུ (du) form or the ལ (su) form, and which they would take would not be immediately knowable based purely on their synchronic forms. For example, two hypothetical words, གས (kas) and གད (kad) would both have the modern forms /kʰɛʔ/. Without the orthography, one could not know whether /kʰɛʔ/ should take the (archaic) oblique form /kʰɛʔ.l.su/ or /kʰɛʔ.l.tu/. One could conveniently brush this problem under the rug by saying that since this form of the oblique is no longer productive, these no longer represent the different forms of some underlying case but are instead entirely part of the phrases in which they are now fossilized.

Similarly, the ablative case nowadays has split based on its usage: the common form གས (nas /nɛʔ/) being used for origin, instrument, and certain adverbial phrases and the form ལས (las /lɛʔ/) used for comparison. However, the ལས form is still, somewhat archaically, sometimes used where གས is, introducing yet another small exception to our neat system.⁷

6 Conclusion

This split between case forms depending on the weight of the preceding syllable pervades Lhasa Tibetan and can be found to a lesser extent in the verbal morphology as well. Changes to grammar (such as the loss of the ergative amongst younger speakers) and the sole usage of the གི (gi) and ལ (la) forms of the genitive and oblique regardless of the preceding syllable may do away with this system, but for now it remains a notable characteristic of Lhasa speech.

The neat system we have laid out here is complicated by certain archaisms deriving from the more complex system of cases found in Classical Tibetan, many of which appear in fossilized phrases in the modern language. However, whether or not these archaisms continue to represent unique forms of their associated cases or if they are now only diachronic rather than synchronic in nature is a question for future research.

7. Nicolas Tournadre and Sangda Dorje, *Manual of Standard Tibetan* (Snow Lion, 2003).