Several language families and a few language isolates are represented in the Himalayas, the world’s greatest massif, running a length of over 3,600 km. The most well-represented language family in this region happens to be the Trans-Himalayan language family, whose very centre of gravity and phylogenetic diversity is situated within the Eastern Himalaya. This most populous language family on our planet in terms of numbers of speakers used to be known as Tibeto-Burman but, in some circles, the family formerly also went by the names “Indo-Chinese” or “Sino-Tibetan”, the latter two labels actually designating empirically unsupported and now obsolete models of language relationship. The study of Trans-Himalayan historical grammar began with Brian Houghton Hodgson in the 1830s, who during this time served at Kathmandu as the British Resident to the Kingdom of Nepal. Periodically, minor studies devoted attention to several of the more salient morphosyntactic phenomena of Trans-Himalayan historical grammar, but Stuart Wolfenden contributed the first major monograph to the subject in the 1920s. Finally, the historical morphosyntax of the Trans-Himalayan language family came to be the focus of numerous linguistic studies from the 1970s onward, and since that time our understanding of the historical grammar of the language family has changed drastically.

As ever more languages out of the hundreds of previously undocumented Trans-Himalayan tongues came to be described and analysed in great detail, it came to be understood that the flamboyant verbal agreement morphology observed in languages such as the Kiranti languages of eastern Nepal and the rGyalrongic languages of southwestern China were neither grammatically innovative nor represented typological flukes, but instead represented the most grammatically conservative languages within the entire language family. Subsequently, cognate inflectional systems or vestiges of cognate conjugational morphology were discovered in most other branches of the language family as well. The geographical centre, as well as the centre of phylogenetic diversity of the Trans-Himalayan language family, was identified as the highland arc of the Eastern Himalaya. Sinitic languages, although representing by far the most populous single branch of the Trans-Himalayan family, were now understood as constituting just one out of many subgroups, not more divergent from other branches than any one of the four dozen other subgroups making up the language family. The various types of epistemic marking systems observed sporadically throughout the region were shown to be secondary innovations, reflecting a great variety of semantically distinct language-specific grammatical categories. Particularly, languages showing the typology of the Loloish or Sinitic type were shown to be innovative in their grammar, having lost much of the original Trans-Himalayan morphosyntax.

Keywords: language families, morphosyntax, Trans-Himalayan language family, Sinitic languages

Himalayan Languages

The Himalayas comprise a vast chain of mountains, hills and plateaus which stretch over a length of more than 3,600 kilometres from the Hazārahjāt highlands of Afghanistan in the west to the Liángshān hills of southern Sìchuān in the east. Numerous language families are represented in this expansive region, including Indo-European, Austroasiatic, Altaic and even Dravidian and Kradai. The
Himalayas also harbour the language isolates Kusunda and Burushaski, which have not been conclusively tied to any known linguistic phylum. Yet most languages in the Himalayas belong to the Trans-Himalayan language family, which in terms of numbers of speakers is the second most populous language family on our planet. The name Trans-Himalayan itself indicates the geographical distribution of the majority of the languages of the family, which straddles the Himalayas from west to east, but which, in terms of the sheer number of languages, is concentrated predominantly in the eastern half of the Himalayas.

Previously, proponents of a certain family tree model had for several decades called the family “Sino-Tibetan”, but as of today no evidence has ever been adduced for that particular phylogenetic model. By contrast, Trans-Himalayan denotes an empirically supported language family model that incorporates only recognised and newly validated subgroups but which remains agnostic regarding phylogenetic relationships for which little or no evidence has been adduced. The new term has been adopted in Chinese literally as 跨喜馬拉雅語系 Kuà-Xīmǎlàyǔ yìxì ‘Trans-Himalayan family’ (Wú, 2015). The language family was first recognised in 1823 by Julius von Klaproth, and was initially referred to by scholars in the British Isles as ‘Tibeto-Burman’, for example, Hodgson (1857), Cust (1878), Forbes (1878) and Houghton (1896). Later, the term ‘Tibeto-Burman’ became ambiguous in usage when proponents of the Indo-Chinese or ‘Sino-Tibetan’ paradigm used the term at variance with its original meaning in a way that excluded the Sinitic languages, whereas in fact the Sinitic languages, such as Cantonese, Mandarin, Hakka and Hokkien, together make up just one of the subgroups within the Trans-Himalayan family, an inventory of which is provided in van Driem (2014).

Trans-Himalayan encompasses well over three hundred languages. Figure 1 shows the geographical distribution of Trans-Himalayan languages, and Figure 2 shows the distribution of the major subgroups of the language family, with each dot representing not a language, but a recognised linguistic subgroup within the family. On the basis of this phylogenetically indicative distribution, it is quite clear that the Trans-Himalayan language family straddles the Himalayas on both sides and that the centre of gravity and diversity lies in the eastern Himalayan region. The difference between the two maps illustrates the results of the historically attested spread within the horizon of recorded history of just a single subgroup, the Sinitic languages, emanating from the eastern end of the original geographical range of the Trans-Himalayan language family.

### Morphosyntactic Conservatism versus Attrition

In terms of their grammatical typology, the Sinitic languages show little if any overt morphology. Chinese mystified linguists such as Wilhelm von Humboldt (1836) with its ‘scheinbare Abwesenheit aller Grammatik’ in the sense that grammatical relations are primarily expressed ‘durch Stellung’. Scott DeLancey (2013a) recently became a proponent of the idea that Sinitic could have arisen as a creole which served as a *lingua franca* over a vast area when the ancestral Tibeto-Burman language spoken by the ancient Chinese was introduced to and adopted by an indigenous Altaic population in what today is northern China. In reality, the idea of an Altaic substrate in Sinitic, and Mandarin in particular, has been propounded for quite some time, for example, Poppe (1965) and Ch’en (1976). In terms of its salient typological features, Mantarō Hashimoto (1976a, 1976b, 1980, 1986) argued that Sinitic could be thought of as an ancient Tibeto-Burman language which underwent pidginisation in the mouths of an Altaic population. Hashimoto adopted William Ballard’s (1979) metaphor, which depicts Sinitic as a mosaic of structural features representing a typological halfway house between Southeast and Northeast Asia. Some have responded critically, for example, Li (1995), whereas others have received the idea favourably, for example, Norman (1982) and Wadley (1996).

Sinitic is not alone as a subgroup within the family in which language contact and population movements appear to have accelerated typological change. In the eastern portion of the Trans-Himalayan geographical range, the Lolo-Burmese and Karenic languages as well as Bái and Tǔjīà have undergone numerous innovations. Even some Himalayan languages show typological similarity to Sinitic or Lolo-Burmese languages. Post (2015) has presented arguments that the morphosyntactic streamlining observed in the Tani languages of Arunachal Pradesh could stem from the area where these languages are spoken having historically been subject to the same upheaval through tribal warfare and perennial migration as the areas of Southeast Asia where the other grammatically more innovative Trans-Himalayan languages are found. The four Trans-Himalayan subgroups that show the greatest conservatism in terms of the retention of the ancestral Trans-Himalayan verbal agreement system are the Kiranti languages of eastern Nepal, the Black Mountain Mónpa language of midwestern Bhutan, the Gongduk language of southeastern Bhutan and the rGyalrongic languages of Sichuān.
The Gongduk and Limbu cases are adduced below. Vestiges of the reconstructible Trans-Himalayan biactantial verbal agreement system are found not just in these four subgroups, but also in other major subgroups such as Dhimlish, Magaric, Nungish, Kachinic also known as Jinghpaw, Newaric, Kukish also known as Mizo-Kuki-Chin, Brahmaputran, West Himalayish and Qiāngic. Much has been written on the antiquity of reconstructible verbal agreement in Trans-Himalayan since James John Bauman’s (1975) doctoral dissertation. Because the interest in the historical morphosyntax of Trans-Himalayan is undergoing a revival in the early 21st century, a list of relevant work is contained in the bibliography so that younger scholars of Trans-Himalayan historical grammar may be less likely to overlook the older literature, not all of which is easy to find and much of which is still highly relevant. The biactantial verbal agreement system of the Gongduk transitive verb and the conjugational morphology of the Phedappe Limbu verb are presented here. Finally, a discussion ensues regarding a newer genre of morphosyntax that has emerged in many Himalayan languages.

The Example of the Gongduk Verb

The Gongduk language forms a phylogenetic subgroup unto itself within the Trans-Himalayan family. The Gongduk language is spoken by a dwindling number of speakers in a remote enclave to the west of the Kurichu in south-central Bhutan, surrounded by Khengkha speaking territory. The native pronunciation of the area is [goŋduk]. The Gongduk personal pronouns are ḍeo ‘I’, ḍīŋ ‘we’, gi ‘you’ (singular), giŋ ‘you’ (plural), gon ‘he, she, it’ and gonma ‘they.’ Separate ergative forms exist for the first person singular ḍeo ‘I’ and the third person pronouns gonde ‘he, she, it’, gonmae ‘they’ and qualifiers such as ḍenpo ‘all’, where the ergative morpheme -e can be readily recognised in each form, though there is no overt ergative morpheme in other personal pronouns. The phonological
transcription used here has been updated with respect to van Driem (2013) and conforms to the forthcoming Gongduk grammar being prepared by Pascal Gerber, Selin Grollmann and myself.

A Gongduk simplex is defined as an inflected non-periphrastic indicative verb form containing person and number agreement affixes but lacking overt aspect markers. In the current analysis, three functional positions or slots are posited to account for the conjugational morphology of the Gongduk verb: one prefixal slot and two suffixal slots. The prefixal slot can be occupied by the negative morpheme \(<\text{mə}>\), and the first suffixal slot can be occupied by the preterite tense morpheme. All agreement suffixes occur in the second suffixal slot.

Table 1 provides the conjugation table of the transitive verbs \(<\text{mə} \sim \text{mɤt}>\) ‘see.’ Each cell in the tables contains four finite forms in the following order: non-preterite affirmative, non-preterite negative, preterite affirmative, preterite negative. The verb \(<\text{mə} \sim \text{mɤt} >\) ‘see’ shows a fixed pattern of stem alternation, as many Gongduk verbs do throughout the agreement paradigm. These complex patterns differ from verb to verb. The stem alternation in the paradigm of the verb \(<\text{mə} \sim \text{mɤt} >\) ‘see’ prima facie resembles a straightforward morphophonologically conditioned pattern of alternation between a preconsonantal and a prevocalic stem, as observed in Limbu. However, Gongduk resembles Dumi more in that all verbs can be grouped into sets or ‘conjugations’ defined by their specific pattern of paradigmatic stem alternation. The conjugation of a verb and its various stems must therefore be specified in the lexicon.

Some Gongduk verbs exhibit tense-motivated stem alternation affecting the stem initial. For example, the verb ‘to give’ has the non-preterite stem \(<\text{pi}>\) and the stem \(<\text{bi}>\) in the preterite, whereas the verb ‘to eat’ shows the non-preterite stem \(<\text{ðə}>\) and the preterite stem \(<\text{θə}>\). In addition to fixed patterns of paradigmatic stem alternation, it is useful to distinguish between the ante-vocalic and the ante-consonantal form of a single stem.

<table>
<thead>
<tr>
<th>Table 1. (&lt;\text{mal-} \sim \text{mɤt}&gt;&gt; ) ‘see’</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2s</td>
</tr>
<tr>
<td>2p</td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>

The Gongduk transitive verb distinguishes 1\(\rightarrow\)2, 1\(\rightarrow\)3, 2s\(\rightarrow\)3, 2p\(\rightarrow\)3, 3\(\rightarrow\)1, 3\(\rightarrow\)2 and 3\(\rightarrow\)3 forms. The notation 1\(\rightarrow\)2 signifies a
The negation and agreement affixes of the Gongduk simplex can be identified as follows:

1. The negative morpheme <nə> occurs as a prefix in all negative indicative forms and is attached immediately to the stem of the verb. The prohibitive <də> in negative imperative forms also occurs in this slot.

2. The preterite morpheme occurs as a suffix attached immediately to the verb stem. Non-preterite tense is expressed as zero. The preterite morpheme exhibits allomorphy with the allomorph <ti ~ di> in 1/2→3 forms (with the allomorphs <t ~ d> appearing in 1/2→3 forms before a vowel), <θə> in 1⇔2 forms, and <ə ~ ɛ ~ ɤ > in all intransitive forms and in transitive forms with a third person agent. The preterite allomorph <ti ~ t ~ di ~ d> in 1→2/3 forms occurs in its voiced form following a nasal and in its unvoiced form after a plosive or vowel. The preterite allomorph <ə ~ ɛ ~ ɤ > occurs as <ɛ> after an open stem with the stem vowel /i/, and as < ɣ > after verb stems with stem vowel < ɣ >, including the verb <mal ~ mət 'see' where the stem vowel alternates, for example, məl-əni 'he/she/they saw him/her/them.' Elsewhere the form < ɟ > of the preterite allomorph occurs. Vowel harmonic rules appear to apply. The morphophonological rule of *vocalis ante vocalem corripitur* holds for the preterite allomorph < ə ~ ɛ ~ ɤ >, which is realised as zero before another vowel—namely, in first person and second person plural intransitive forms—and for the allomorphs <ti ~ di> in 1→2/3 forms, which are reduced to <t ~ d> before a vowel.

3. The 1⇔2 morpheme < ne ~ ni ~ no > marks a transitive relationship between a first and a second person actant. The form <ni> is a vowel harmonic form of the morpheme when the preceding vowel is /i/, /u/ or /i/. The form <ne> occurs when the preceding vowel is not a closed vowel, for example, after /ɛl/, /ɔ/ or /o/. The 1⇔2 morpheme has a regular allomorph <nə> in the negative.

4. The 1→3 morpheme < uŋi ~ ore > indexes a transitive relationship between a first person agent and a third person patient. The form <uŋi> is a vowel harmonic form of the morpheme when the preceding vowel is /i/, /u/ or /i/. In allegro speech, the segment /ŋ/ can become realised as nasality and the suffix consequently becomes totally vocalic in nature, i.e., [u]. The form <ore> occurs when the preceding vowel is not a closed vowel, e.g., after /ɛl/, /ɔ/ or /o/. The 1→3 morpheme has a regular allomorph <uŋ > in the negative. Glottal hiatus is observed between the ending and the stem vowel of the open stem verb ‘to give’, for example, pi-ʊɛni 'l/we shall give it to him/her/them', but it is not observed in the other transitive open stem verb which occurs in the corpus, ‘to eat’, for example, ɔŋne 'l/we (shall) eat [it].'

5. The 2s→3 morpheme < ni ~ ne > expresses a transitive relationship between a second singular agent and a third person patient. The form <ni> is a vowel harmonic form of the morpheme when the preceding vowel is /i/, /u/ or /i/. The form <ne> occurs when the preceding vowel is not a closed vowel, for example, after /ɛl/, /ɔ/ or /o/. The same morpheme indexes a second singular subject in intransitive verbs, but this analysis will be limited to the transitive agreement system.

6. The 3→1/2 morpheme < Ti ~ Te ~ Tə > expresses a transitive relationship between a third person agent and a first or second person patient. The same morpheme indexes a third person subject in non-preterite intransitive verbs. The vowel harmonic allomorph < Ti > occurs after the vowels /i/, /u/ or /i/. The vowel harmonic allomorph < Te > occurs after non-closed vowels /el/, /ɔ/ or /o/. The archiphoneme /T/ is /t/ after a plosive, /d/ after a nasal, /r/ after a vowel and /n/ in the preterite. The 3→1/2 morpheme has a regular allomorph < ni ~ ne > in the preterite and < nə > in the negative preterite. The allomorphic pattern of the 3→1/2 morpheme within the preterite is reminiscent of the allomorphic pattern of the 1⇔2 morpheme < ne ~ ni ~ no >.

7. The 2p→3 morpheme < uri ~ ore > indexes a transitive relationship between a second plural agent and a third person patient. The form <uri> is a vowel harmonic form of the morpheme when the preceding vowel is /i/, /u/ or /i/, but also, it appears, following the stem <məl> of the verb <mal ~ mət 'see', where the stem vowel /ə/ alternates with /ɛ/, for example, mətururi 'you (plural) saw him/her/them.' The form <ore> occurs when the preceding vowel is not a closed vowel, for example, after /ɛl/, /ɔ/ or /o/. The 2p→3 morpheme has an allomorph < rɨ ~ re > immediately following
the vowel of an open stem verb.

Three additional morphemes occur in the Gongduk intransitive paradigm, i.e., the non-preterite first person subject morpheme \(< \gamma\mathbf{g}i \sim \gamma\mathbf{n}>\), preterite first person subject morpheme \(< \gamma\mathbf{n}\mathbf{i} \sim \gamma\mathbf{n}>\) and the second plural subject morpheme \(< \text{iri}>\).

### The Example of the Phedappe Limbu Verb

Limbu is a Kiranti language native to eastern Nepal and the western fringe of Sikkim. Phedappe Limbu is the phonologically most conservative language within the Limbu branch of Kiranti, the internal phylogeny of which is outlined in van Driem (2001, p. 615). The morphological analysis of the Limbu simplex verb presented in van Driem (1987) was revised in van Driem (1997, 1999).

Simplicia are non-periphrastic finite indicative verb forms, from which the various Limbu periphrastic tense forms, gerunds and participles are derived. Limbu distinguishes eleven pronominal categories, namely first, second and third person, singular, dual and plural number, and there is an inclusive versus exclusive distinction in the first person dual and plural. The Limbu transitive verb shows agreement with both agent (transitive subject) and patient (transitive object or beneficiary), and the transitive paradigm distinguishes 44 different forms, as shown in tables 2 and 3. Limbu intransitive and reflexive verbs agree with the subject (intransitive or reflexive subject), and the intransitive and reflexive paradigm distinguishes eleven different forms, as shown in tables 4 and 5.

The Kiranti languages of eastern Nepal and the rGyalrongic languages of Sichuăn show richer prefixal morphology than do other Trans-Himalayan languages. Kiranti verbs exhibit a larger repertoire of agreement morphemes and distinguish more slots or functional positions in their affixal strings than do verbs in Gongduk, Jinghpaw, Magaric or Black Mountain Monpa, although rGyalrongic verbs can be just as complex in their agreement morphology as a Kiranti verb. The first functional slot pf1 can be occupied by the first person prefix \(< \mathbf{a}>\) (1), which precedes other prefixes in the affixal string. The second person prefix \(< \mathbf{k}\mathbf{e}>\) (2) indexes second person in forms in which second person involvement is not indexed by some *portemanteau*. The non-singular agent/subject morpheme \(< \mathbf{m}\mathbf{e} \sim \mathbf{m}>\) (3nsAS) specifically marks the involvement of a third person non-singular agent or subject. Its abbreviated allomorph \(< \mathbf{m}>\) occurs between a preceding prefix and the root of the verb, but the full form occurs between a preceding prefix and a following negative morpheme. Involvement of a third person singular agent or subject (3sAS) is marked by zero. Negation in the verb is marked by a single ‘discontinuous morpheme’ or by what Claude Hagège (1986, p. 26) called a *simulfixe*. A simulfix consists simply of two or more elements which are affixed simultaneously at different locations in the affixal string. The labels (NEG₁), (NEG₂) and (NEG₃) with different subscripts can be used to indicate the elements of the Limbu negative simulfix. A Limbu negative may contain from one to three negative morphs, the occurrence of which is paradigmatically determined and one of which is always a prefix. The second and third negative morphemes occur in suffixal slots sf6 and sf10. The prefixal string of the Limbu simplex is structured as follows:

### The Phedappe Limbu prefixal string

\[
\begin{array}{lcccc}
\text{pf1} & \text{pf2} & \text{pf3} & \text{pf4} \\
<\mathbf{m}\mathbf{e} \sim \mathbf{m}> & 3\text{nsAS} & <\mathbf{m}\mathbf{e} \sim \mathbf{n} \sim \mathbf{m}\mathbf{e}\mathbf{n}> & \text{NEG₁} \\
\text{<a>} & \text{<ke> } & \text{2} & \text{3sAS} \\
\text{1} & \text{2} & \text{ } & \text{ } \\
\text{Ø} & \text{ } & \text{ } & \text{ }
\end{array}
\]

As in all Trans-Himalayan languages retaining reflexes of the older biactantial agreement system, the longer suffixal string in the Phedappe Limbu verb hosts a larger repertoire of agreement morphemes than the prefixal string. The first suffixal slot (sf1) can be
occupied by the reflexive morpheme or the 1 → 2 portemanteau morpheme < nɛ ~ n > (1 → 2). The reflexive morpheme < siŋ ~ nɛ ~ n > (REF) shows the allomorph < nɛ ~ n > in dual forms and the allomorph < siŋ > in singular and plural forms, and the morpheme denotes reciprocal or reflexive meaning. The 1 → 2 portemanteau morpheme < nɛ ~ n > (1 → 2) indexes the transitive relationship between a first person agent and a second person patient.

In reflexive forms, the dual reflexive allomorphs < nɛ ~ n > always occur as part of the sequence < nɛtchi >, whereby the element < tchi > is the regular allomorph < tchi > of the dual patient/subject morpheme < si ~ tchi > (dPS) in suffixal slot sf3. The allomorph < tchi > occurs regularly following the preterite morpheme < ŋ > (PT), the dual reflexive allomorph < nɛ ~ n > (REF) and the 1 → 2 portemanteau < nɛ ~ n > (1 → 2). The dual reflexive preterite sequence < n-ɛtchi > (REF-PT-dPS) is homophonous with the corresponding non-preterite sequence < nɛ-Ø-tchi > (REF-NPT-dPS). Consistently throughout the paradigm, cases of preterite/non-preterite homophony are satisfactorily accounted for by adjacency of a vocalic morpheme to the preterite morpheme < ŋ >, which elides. The dual reflexive allomorphs < nɛ ~ n > are homophonous with the allomorphs of the 1 → 2 portemanteau morpheme < nɛ ~ n >, which occurs in the 1s → 2s ending < nɛ >, the 1s → 2d ending < nɛtchiq >, the 1s → 2p ending < niŋ > and the 1ns → 2 ending < nɛtchige >.

The second suffixal slot sf2 is the tense slot. Non-preterite time is marked by zero, and preterite time is marked by the suffix < ŋ > with a zero allomorph before another vowel. The non-preterite zero suffix does not occur in forms in which tense is indexed by the non-preterite first person singular patient/subject morpheme < ŋ > (1sNS/PT). The old dual agent morpheme < s ~ tch > (d) is the sole potential occupant of suffixal slot sf3. The morpheme < u > (3P) marks the involvement of a third person patient. The morpheme < p > is a preterite third person patient morpheme (3P/PT). The morpheme < p > (3P/PT) occurs in the morpheme sequence < paŋ > in negative preterite 1s → 3 forms, for example, mɛbi:-baŋ-si-ŋ ‘I didn’t give it to them’ (1s → 3ns/PT), cf. mɛbi:-Pɛ-n-chi-n ‘I shan’t give it to them’ (1s → 3ns/NPT), but the suffix < paŋ > is optional in intransitive negative preterite forms with a first person singular subject. This leads to doubles, such as yaʔ-mɛ恩-la:k-paŋ (paddy.dance-NEG-paddy.dance-3P-PT-1s) ‘I did not partake in the rice harvest dance’ and yaʔ-mɛ- ра:тн-аŋ-иɛ (paddy.dance-NEG-paddy.dance-1s-NEG) with the same basic meaning. Where such doubles exist, difference in form expresses difference in meaning, even when the nuance is subtle. The suffix < paŋ > appears to have originally been limited to the transitive paradigm and was secondarily absorbed into the intransitive paradigm, which corresponds neatly with the difference in meaning between negative preterite doubles with a first singular subject.

Negative preterite first singular intransitive forms in < paŋ > coexist with the ‘regular’ intransitive forms. The intransitive forms kɛray ‘I have arrived’ and khi:raŋ ‘I am [i.e., have become] tired’ are negated either as mɛgɛrαŋɛn ‘I haven’t arrived’ and mɛkhi:raŋɛn ‘I am not tired’ or as mɛγɛpɛŋ and mɛγkhi.ppαŋ, respectively. The difference in meaning is that the ‘regular’ first singular intransitive affixation in < mɛ-Σŋ-nɛn > is simply the negative alternative to its affirmative preterite counterpart in < aŋ >, for example, kɛray: mɛɡɛrαŋɛnɛni: ‘Have I arrived or haven’t I?’ Note that the symbol Σ denotes the verb root. The use of the suffix < paŋ > in negative intransitive forms, however, adds a semantic dimension to the utterance. The added implication is that something hasn’t happened that was supposed to happen or was expected to happen. For example, as an answer to a question one can either respond with mɛłхаγɛнɛn ‘I didn’t come up’, the regular intransitive negative of thαŋ ‘I have come up’, or with the more abrupt mɛndhaŋɛн. So, the tensed third person patient morpheme < p > (3P/PT) may occur in negated first person intransitive forms, whereby third person patient agreement with the logical proposition denoted by the verb and not enacted by the first person singular agent, for example, mɛn-dhaŋ-bay in the sense ‘I didn’t do it, viz. come up.’ The latter form with the originally transitive suffix < paŋ > can be construed as rude because the implication is: I was supposed to have come up, but I didn’t. The regular intransitive form mɛłхаγɛнɛน by contrast is a simple negation and is therefore more polite, that is, ‘I didn’t come up’ or, by implication, ‘I couldn’t come up.’

The meaning of the forms in < paŋ > has something to do with the greater volitionality or agency of this originally transitive suffix. In combination with the lexical meaning of certain verbs the form in < paŋ > is the more obvious choice, for example, mɛnchibay ‘I’m not dead, I didn’t die’, suitable when the speaker has just survived a life threatening situation or when the speaker, presumed dead, turns out not to have died. This is why the form mɛšyaŋɛн ‘I’m not dead’ is an odd, albeit acceptable, utterance in modern Limbu, for it is as if the speaker is making the remark out of the blue. There are reportedly few conceivable situations in which mɛšyaŋɛн would be a relevant utterance.
The non-preterite first person singular suffix `<ʔɛ>` (1s/NPT) indexes the involvement of a first person singular actant in non-preterite time. Singular number of a first person actant is indexed by the first person singular morpheme `<aŋ ~ ŋ>` (1s) or the non-preterite first person singular morpheme `<ʔɛ>` (1s/NPT). The plural agent suffix `<m>` (pA) in suffixal slot sf5 follows the third person patient morpheme `<u>` and precedes the non-singular morpheme `<si ~ s ~ tchi ~ tch>` in suffixal slot sf7. In the same functional position, the portemanteau morpheme `<mʔna>` (1peAS/PT) indexes a first person plural exclusive subject or agent in preterite time. In suffixal slot sf5, the plural patient/subject morpheme `<i>` (pPS) indexes plural number of a first or second person patient in 3→1pi, 3→1pe, 1s→2p and 3→2p forms, and plural number of first or second person subject in intransitive forms. The morpheme also occurs in intransitive adhortative forms with a first person plural inclusive subject.
The transitive paradigm does not distinguish a dual from a plural third person patient, and non-singular number of a third person patient is indexed by the non-singular morpheme < si ~ tchi > (ns) in suffixal slot sf7. The morpheme also indexes non-singular number of agent in 1nse→2 forms, and the morpheme shows the regular allomorph < tchi > after either the preterite tense suffix < ɛ > or the 1→2 portemanteau < nɛ >. Singular number is marked by zero. The copy slot, suffixal slot sf8, contains copy morphemes of the first person singular copy morpheme < η > (1s) and the plural agent copy morpheme < m > (pA), both of which occur primordially in suffixal slot sf5. Copy morphemes can be said to be paradigmatically conditioned or triggered by the occurrence of a limited set of other morphemes, including mainly the non-singular suffix < si ~ tchi > (ns). Whereas the plural agent copy morpheme < m > (pA) echoes the homophonous agentive morpheme, the first person singular copy < η > (1s) reiterates the notion of first singular agent indexed by either the first singular morpheme < η > (1s) in suffixal slot sf5 or by the 1→2 portemanteau morpheme < nɛ ~ n > (1→2) in suffixal slot sf1.

The third element of the negative simulfix < n > (NEG3) immediately precedes the slot containing the non-singular number morpheme < si > (ns), the presence of which triggers the occurrence of this third element in negative forms. In suffixal slot sf9, the exclusive suffix < ge ~ be > (e) indexes first person involvement and the exclusion of second person, whilst the inclusion of second person in first person forms is formally zero-marked.

The transitive paradigm does not distinguish a dual from a plural third person patient, and non-singular number of a third person patient is indexed by the non-singular morpheme < si ~ tchi > (ns) in suffixal slot sf7. The morpheme also indexes non-singular number of agent in 1nse→2 forms, and the morpheme shows the regular allomorph < tchi > after either the preterite tense suffix < ɛ > or the 1→2 portemanteau < nɛ >. Singular number is marked by zero. The copy slot, suffixal slot sf8, contains copy morphemes of the first person singular copy morpheme < η > (1s) and the plural agent copy morpheme < m > (pA), both of which occur primordially in suffixal slot sf5. Copy morphemes can be said to be paradigmatically conditioned or triggered by the occurrence of a limited set of other morphemes, including mainly the non-singular suffix < si ~ tchi > (ns). Whereas the plural agent copy morpheme < m > (pA) echoes the homophonous agentive morpheme, the first person singular copy < η > (1s) reiterates the notion of first singular agent indexed by either the first singular morpheme < η > (1s) in suffixal slot sf5 or by the 1→2 portemanteau morpheme < nɛ ~ n > (1→2) in suffixal slot sf1.

The third element of the negative simulfix < n > (NEG3) immediately precedes the slot containing the non-singular number morpheme < si > (ns), the presence of which triggers the occurrence of this third element in negative forms. In suffixal slot sf9, the exclusive suffix < ge ~ be > (e) indexes first person involvement and the exclusion of second person, whilst the inclusion of second person in first person forms is formally zero-marked.
Table 2. Endings of the Phedappe Limbu affirmative transitive verb. (In each cell, the preterite form is listed below the corresponding non-preterite form.)

<table>
<thead>
<tr>
<th></th>
<th>1s</th>
<th>1di</th>
<th>1de</th>
<th>1pi</th>
<th>1pe</th>
<th>2s</th>
<th>2d</th>
<th>2p</th>
<th>3s</th>
<th>3d</th>
<th>3p</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Σ-ae</td>
<td>Σ-ætæg</td>
<td>Σ-ætig</td>
<td>Σ-a</td>
<td>Σ-apig</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1di</td>
<td>Σ-a</td>
<td>Σ-aætæ</td>
<td>Σ-aætæchæ</td>
<td>Σ-aætæchæ</td>
<td>Σ-aætæchæ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1de</td>
<td>Σ-aætæ</td>
<td>Σ-aætæge</td>
<td>Σ-aætæge</td>
<td>Σ-aætæge</td>
<td>Σ-aætæge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1pe</td>
<td>Σ-aætæge</td>
<td>Σ-aætæge</td>
<td>Σ-aætæge</td>
<td>Σ-aætæge</td>
<td>Σ-aætæge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>2s</td>
<td>ke-Σ-a</td>
<td>ke-Σ-an</td>
<td>ke-Σ-en</td>
<td>ke-Σ-ætæ</td>
<td>ke-Σ-ætæchæ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3d</td>
<td>ke-Σ-ætæ</td>
<td>ke-Σ-ætæge</td>
<td>ke-Σ-ætæge</td>
<td>ke-Σ-ætæge</td>
<td>ke-Σ-ætæge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>3p</td>
<td>ke-Σ-ætæge</td>
<td>ke-Σ-ætæge</td>
<td>ke-Σ-ætæge</td>
<td>ke-Σ-ætæge</td>
<td>ke-Σ-ætæge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1s</td>
<td>Σ-æn</td>
<td>Σ-æntsæ</td>
<td>Σ-æntsæchæ</td>
<td>Σ-æntsæchæ</td>
<td>Σ-æntsæchæ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1di</td>
<td>Σ-æntsæchæ</td>
<td>Σ-æntsæchæ</td>
<td>Σ-æntsæchæ</td>
<td>Σ-æntsæchæ</td>
<td>Σ-æntsæchæ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1pe</td>
<td>Σ-æntsæchæ</td>
<td>Σ-æntsæchæ</td>
<td>Σ-æntsæchæ</td>
<td>Σ-æntsæchæ</td>
<td>Σ-æntsæchæ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>2s</td>
<td>ke-Σ-æntsæchæ</td>
<td>ke-Σ-æntsæchæ</td>
<td>ke-Σ-æntsæchæ</td>
<td>ke-Σ-æntsæchæ</td>
<td>ke-Σ-æntsæchæ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3d</td>
<td>ke-Σ-æntsæchæ</td>
<td>ke-Σ-æntsæchæ</td>
<td>ke-Σ-æntsæchæ</td>
<td>ke-Σ-æntsæchæ</td>
<td>ke-Σ-æntsæchæ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3p</td>
<td>ke-Σ-æntsæchæ</td>
<td>ke-Σ-æntsæchæ</td>
<td>ke-Σ-æntsæchæ</td>
<td>ke-Σ-æntsæchæ</td>
<td>ke-Σ-æntsæchæ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Endings of the Phedappe Limbu negative transitive verb. (In each cell, the preterite form is listed below the corresponding non-preterite form.)

<table>
<thead>
<tr>
<th></th>
<th>1s</th>
<th>1di</th>
<th>1de</th>
<th>1pi</th>
<th>1pe</th>
<th>2s</th>
<th>2d</th>
<th>2p</th>
<th>3s</th>
<th>3d</th>
<th>3p</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>mo-Σ-æn</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
</tr>
<tr>
<td>1di</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
</tr>
<tr>
<td>1de</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
</tr>
<tr>
<td>1pe</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
<td>mo-Σ-œntæchæ</td>
</tr>
</tbody>
</table>
Endings of the Phedappe Limbu intransitive and reflexive verb. (In each cell, the negative form is listed below the corresponding affirmative form.)

<table>
<thead>
<tr>
<th></th>
<th>INTRANSITIVE</th>
<th></th>
<th>REFLEXIVE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>NPT</td>
<td>PT</td>
<td>NPT</td>
</tr>
<tr>
<td>1s</td>
<td>Σ-te</td>
<td>Σ-an</td>
<td>me-Σ-ten</td>
<td>me-Σ-anən</td>
</tr>
<tr>
<td>1di</td>
<td>a-Σ-si</td>
<td>a-Σ-etchi</td>
<td>an-Σ-sin</td>
<td>an-Σ-etchin</td>
</tr>
<tr>
<td>1de</td>
<td>Σ-sige</td>
<td>Σ-etchige</td>
<td>me-Σ-sigen</td>
<td>me-Σ-etchigen</td>
</tr>
<tr>
<td>1pe</td>
<td>Σ-ige</td>
<td>Σ-m?na</td>
<td>men-Σ-m?na</td>
<td></td>
</tr>
<tr>
<td>2s</td>
<td>ke-Σ</td>
<td>ke-Σ-e</td>
<td>ken-Σ-nen</td>
<td>ken-Σ-en</td>
</tr>
<tr>
<td>2d</td>
<td>ke-Σ-si</td>
<td>ke-Σ-etchi</td>
<td>ken-Σ-sin</td>
<td>ken-Σ-etchin</td>
</tr>
<tr>
<td>2p</td>
<td>ke-Σ-i</td>
<td>ke-Σ-i</td>
<td>ken-Σ-in</td>
<td></td>
</tr>
<tr>
<td>3s</td>
<td>Σ</td>
<td>Σ-e</td>
<td>me-Σ-nen</td>
<td>me-Σ-en</td>
</tr>
<tr>
<td>3d</td>
<td>Σ-si</td>
<td>Σ-etchi</td>
<td>me-Σ-sin</td>
<td>me-Σ-etchin</td>
</tr>
<tr>
<td>3p</td>
<td>me-Σ</td>
<td>me-Σ-e</td>
<td>men-Σ-nen</td>
<td>men-Σ-en</td>
</tr>
</tbody>
</table>

Epistemic and Evidential Systems

Vestiges of old Trans-Himalayan morphosyntax and biactantial conjugational morphology have been retained in many of the linguistic subgroups which make up the language family—namely Kiranti, Dhimalish, Magaric, Nungish, Kachinic, Newaric, Kukish, Brahmaputran, West Himalayish, Gongduk, rGyalrongic, Black Mountain Monpa and Qiángic. However, many branches preserve little trace of the original verbal agreement system. After the erosion of the older morphosyntactic system of encoding actants in the verb, new systems have arisen in many Trans-Himalayan languages. Many of these new systems encode information in the verb which has been termed epistemic or evidential, and such systems have arisen even in some languages of some of the subgroups just mentioned. DeLancey (1992) showed that these systems, although similar, are not cognate and appear all to have arisen independently.

In Bhutan, for example, the Dzongkha verb distinguishes a witnessed past from an inferred past. The interaction of these two grammatical categories with the lexical meanings of individual verbs leads to some fascinating results. In one study (van Driem, 2007), a number of utterances are adduced which a native speaker of Dzongkha would reject as ungrammatical if the sentences are presented without context. As soon as each of the highly specific contexts is explained in which these sentences were actually uttered, then the “ungrammatical” verb forms are suddenly no longer rejected, but turn out instead to be the required verb form in each of the
given instances. Only careful observation can capture such cases, which can help to unravel the semantics of these grammatical categories and illustrate them insightfully.

Another type of epistemic distinction is observed in Newar. The terms conjunct and disjunct were famously introduced by Austin Hale (1980) based on the phenomenon that, in Newar reported speech, the verb of a subordinate clause is conjunct when the subjects of main and subordinate clause are coreferential, and disjunct when they are not. In Newar simple sentences, the conjunct is the form usually used with a first person subject in statements and with a second person subject in questions, and the disjunct is the form usually used with second and third person subject in statements and with first and third person subjects in the interrogative. The difference in how these two categories are distributed in normal speech across the three person categories is interesting. It is salubrious to recall that epistemic categories in the Tibetan verbal system were once misinterpreted as imperfect person agreement systems, for it lies in the very nature of the grammaticational meanings conveyed in evidential systems that they interact unevenly with the person categories because they encode information about the epistemic stance of the subject of the sentence with respect to the situation depicted by the verb in the sentence. Because the meanings of the Newar categories had been inadequately described, with the help of Newar friends additional examples of conjunct and disjunct in Kathmandu Newar were adduced to yield insight into the language-specific meanings of these two grammatical categories because the utterances had been attested in highly specific contexts which lay more or less precisely at the knife-edge of the semantic distinction (van Driem, 2001, pp. 761–766).

Each of these so-called epistemic categories is specific in meaning to the language in which they happen to occur. The category called the ajan bhāt ‘unknown past’ in Nepali, an Indo-European language of the Indo-Aryan branch, but which Michailovsky (1996) has more aptly termed the Nepali inferential, has a demonstrably different language-specific meaning that differs subtly from that of the Dzongkha inferred past, which in Dzongkha grammar takes part in an equipollent morphosyntactic opposition between inferred vs. witnessed past. Examples illustrating the fine distinction in meaning between the Nepali inferential past and the Dzongkha inferred past have been adduced in van Driem (2001, pp. 644–646). The fact that verbal systems have arisen which encode the epistemic stance of either the subject or the speaker, depending on the language and its grammar, has led to a trend whereby typologists coin terms such as ‘mirativity’ or ‘epistemic’ as cover terms for such phenomena. Inevitably, some linguists have used such labels in a Platonic essentialist fashion, whereas others have been careful to word their formulations in such a way that only the reality of language-specific grammatical categories of meaning is acknowledged and that formulations which implicitly or explicitly acknowledge or assume the reality of transcendental cross-linguistic Platonic essences such as ‘aspect’ or ‘mirativity’ are eschewed. There are also language typologists who abjure or publicly deny the reality of Platonic essentialist categories in grammar but whose writings cannot be made sense of unless the reality of such cross-linguistic essences is assumed.

It is on this very topic or non-topic of ‘mirativity’ that a tempest in a teapot has broken loose amongst linguists, many of whom work on Trans-Himalayan languages. The history of linguistics demonstrates that typologists’ coinages essentially give rise to an ad hoc religious belief in the reality of the language-independent, transcendant or cross-linguistic categories thus labelled. For the history of ‘aspect’, see van Driem (2001, pp. 649–657). The storm was perhaps unleashed by the arguably innocent use of the term ‘mirativity’ by DeLancey (1997, 2002). This led to a protest by Hill (2012, 2013), who deemed the usage of the term in the essays in question to be anything but innocent. This critique in turn provoked an impenitent response from DeLancey (2012) as well as jeering support from linguists adopting a notably explicit essentialist stance (Aikhenvald, 2012; Hengeveld & Olbertz, 2012), whereas other more philosophical linguists seized upon the occasion to reflect in insightful ways on what distinguishes and perhaps sometimes unites categories for which language typologists coin labels or to which linguistic descriptivists affix familiar labels by force of analogy or on the basis of presumed similarity to known categories (Friedman, 2012; Hyslop, 2014).

In the realm of morphosyntax, it is noteworthy that the field of Trans-Himalayan linguistics has featured such controversy, both with regard to the biactantial agreement systems in the verb, the antiquity of which the so-called ‘Sino-Tibetanists’ were once wont to dispute, as well as with regard to the morphosyntactic systems which have developed in Himalayan languages more recently, encoding different types of evidential or epistemic meaning. The topographical diversity of the Himalayas is dramatically reflected in the diversity of the morphosyntactic systems attested in the languages spoken there.

References


Find this resource:


Language and Linguistics, 1(2), 211–232.


