# THE TYPOLOGICAL POSITION OF THE LANGUAGES OF NEPAL: MORPHOSYNTAX

*Mark Donohue*[a] *& Siva Kalyan*[b]

[a]The Living Tongues Institute for Endangered Languages, USA
[b]The University of Queensland and The Australian National University, Australia
Corresponding author: mhdonohue@gmail.com

*We examine the typological position of the languages of Nepal among the languages of the world, considering a broad sample of morphosyntactic features. Following a computational analysis, we find that: most of the languages of Nepal occupy a fringe position; some languages of Nepal lie outside the range of variation of other South Asian languages; and the languages of Nepal show correlations among morphosyntactic features that are not correlated globally.*

*Keywords*: Typology, morphosyntax, Nepal, computational methods, language families

## 1. Introduction

In this article we examine the typological position of the languages of Nepal, as judged from a database of morphosyntactic features that covers 79 languages from Nepal (see Appendix) or its immediate surrounds, 276 languages from other countries in South Asia, and 2734 languages from outside this region. This article complements, and presents an updated perspective on, the typology of the languages of Nepal presented earlier in Donohue (2019). The current findings reveal that: most of the languages of Nepal, along with most other languages in South Asia, occupy a fringe position in the range of global morphosyntactic typological variation; some languages of Nepal lie outside the range of variation of the other South Asian languages; and the typologies of the languages of Nepal show correlations among globally uncorrelated dimensions of morphosyntactic variation.

## 2. Global variation in morphosyntax

A recent study (Kalyan & Donohue, 2023) examined a database of 3089 languages, and determined that the morphosyntactic variation among these languages can best be accounted for in terms of four dimensions of variation. Table 1 shows those four dimensions in summary, along with a sample of features that positively correlate with each dimension. The first and fourth dimension can largely be defined in terms of word order features, with Dimension 1 tracking the difference between languages with OV or VO order at the clause level; additionally, languages low on Dimension 1 tend to have case marking, while languages high on Dimension 1 tend towards prefixal agreement on the verb (this is described in more detail in Kalyan & Donohue, 2023). The difference between subject-initial and verb-initial languages is tracked on Dimension 4, in addition to a number of NP-internal word order features. The second and third dimensions correspond to morphological features, with languages high on Dimension 2 showing increasingly elaborate verbs, and languages high on Dimension 3 more likely to have gender systems and plural marking on nouns, as well as agreement by suffix, rather than prefix.

These four dimensions are, by virtue of how they are produced, uncorrelated with each other at the global level. However, correlations may emerge at the local level. In Table 2 we see that the correlation between Dimensions 1 and 2, and between Dimensions 3 and 4, is insignificant at the global level (values are $r^2$ correlations). If we examine just the languages of South Asia we find the same lack of correlation between Dimensions 1 and 2, but slight correlation between dimensions 3 and 4. Restricting ourselves purely to the languages of Nepal we observe that there are strong and significant correlations between Dimensions 1 and 2, and correlations that are nearly as strong between Dimensions 3 and 4. Not only is it surprising that two globally uncorrelated

pairs of dimensions are not in fact uncorrelated in a Nepalese context; it is also surprising that the correlations we find are not between the two word order dimensions or between the two morphology dimensions, but between one word order dimension and one morphology dimension in each case. As we will see in Section 3, these correlations reflect the Tibeto-Burman vs. non-Tibeto-Burman genealogical divide among the languages of Nepal.

Table 1. *The major dimensions of global morphosyntactic variation*

| | Feature types | High values |
|---|---|---|
| 1. | Clausal word order | VO order<br>No case<br>Prefixed agreement |
| 2. | Verbal morphology | Verb agreement<br>Modality affixes<br>Applicatives<br>Possession prefix |
| 3. | Nominal morphology | Plural marking<br>Gender<br>Suffixed verbal<br>   agreement |
| 4. | Clausal word order | VS order<br>Modifier-noun<br>Relative pronoun |

Table 2*: Correlations between dimensions*

| | Global | Sth Asia | Nepal |
|---|---|---|---|
| 1 & 2 | 0.03 | 0.02 | 0.61 |
| 3 & 4 | 0.02 | 0.12 | 0.53 |

Examples illustrating the co-occurrence of features associated with Dimension 1 and Dimension 2 in Nepalese languages can be seen in (1), showing data from Limbu (van Driem, 1987, p. 315) and (2) (Nepali). Limbu, which is high (for a Nepalese language) on Dimension 1, and high on Dimension 2, displays prefixal agreement and possessive marking, while Nepali, low on Dimension 1 and low (for a Nepalese language) on Dimension 2 has suffixal agreement on the verb, and no agreement indexing. Both languages have case marking systems, as is

expected for languages on the lower end of Dimension 1.

Limbu

(1) kɛ-ndzum-in-nu     abhɛlle
your-friend-ABS-COM     when
kɛ-dum-ɛ-tchi?
2-meet-PRETERITE-DUAL
'When did you meet your friend?'

Nepali

(2) रामले उसको साथीलाई कहिले भेटे
Rām-le    us-ko     sāthi-lāi
Ram-ERG   3SG-POSS   friend-DAT
kahile   bheṭ-e
when    meet-3SG.PAST
'When did Ram meet his friend?'

These correlations can be seen visually in Figures 1 and 2. In Figure 1 the languages of Nepal (black) and South Asia (grey) are plotted on Dimensions 1 and 2; while the graph as a whole shows no correlation between the two dimensions, it is clear that within Nepal, an increase or decrease on one of these dimensions corresponds to an increase or decrease on the other. (Figures 1–3 mark the $x$ and $y$ axes in increments of 5, such that the languages shown in Figure 1 run from approximately –10 to +11 on the $x$-axis, and from –14 to +12 on the $y$-axis.)

Figure 1. *Dimensions 1 and 2*
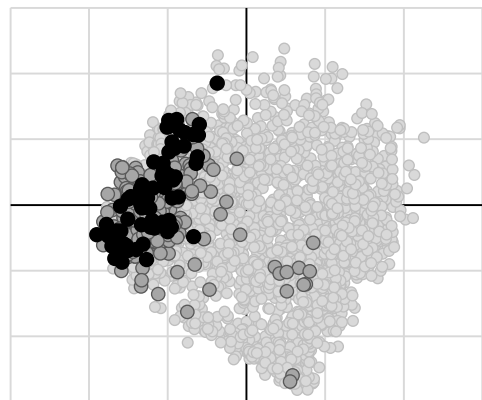


In Figure 1 we can also see that the languages of Nepal, and indeed most of South Asia, occupy a position at the extreme low end of Dimension 1,

and for the most part in the centre of Dimension 2. Moreover, as the first dimension increases, among the languages of Nepal the second dimensions also increases. For the languages of Nepal, it is then reasonable to treat dimensions 1 and 2 as a single combined variable.

The Nepalese language at the high end of Dimension 2, a significant outlier from the other languages of Nepal and South Asia, is Chhatthare Limbu (Tumbahang, 2007), found at (–1.8, +9.2)), which has a significantly more head-marking profile than the other languages in South Asia. At the bottom left of the range of Nepalese languages, and at the left edge of the global set of languages, is Chantyal (–9.5, –2.25), which extends slightly lower on Dimension 1 than do Kurtöp and Lhokpu (from Bhutan) and Jirel (from Nepal). The outlier at the bottom right of the group of Nepalese languages, though firmly embedded within the range of global variation, is Kusunda (–3.4, –2.4), which is typologically more similar to various Kuki-Chin languages (Bawm, Mizo) on the periphery of the South Asian area than to other languages of its immediate area in Nepal. We note that the South Asian languages from outside Nepal which are geographically and typologically peripheral to South Asia form a cluster at (+4, –5): this includes the Khasian languages of Meghalaya, and the Austro-Asiatic languages of the Nicobars. The South Asian languages at the bottom of Figure 1 at (+2.75, -13.5) are unassimilated Tai languages from Assam and Nagaland.

In Figure 2 the languages are plotted on dimensions 3 and 4. As indicated in Table 1, the high end of Dimension 4, the *y*-axis in Figure 2, is occupied by languages with verb-initial clausal word orders; there are no languages in Nepal with this profile. Thus, the languages of Nepal occupy a central position in this figure, and are completely missing from the top left of the plot (the only South Asian languages in this corner of the plot are the verb-initial Austro-Asiatic languages of the Nicobar islands). The bottom right edge of the plot is likewise devoid of any languages from South Asia. The languages of Nepal exist in a space such that these two dimensions correlate, again indicating (as we saw

with Dimensions 1 and 2) that we can combine Dimensions 3 and 4 together as a single variable.

The cooccurrence of Dimensions 3 and 4 are shown in (3) and (4), compared to (5) and (6). The Nepali examples in (3) and (4) illustrate the cooccurrence of prenominal adjectives, agreement suffixes, and gender, typical of a language high on dimensions 3 and 4, while the Tibetan examples in (5) and (6) show postnominal adjectives and no gender or agreement system, as is expected of a language more in the middle of the range of these two dimensions. Within Nepal we do *not* tend to find languages with (for example) postnominal adjectives with gender systems, or languages with prenominal adjectives and yet lacking suffixal agreement. This sort of language is attested elsewhere in the world (eg., the Bantu languages of Africa or the Romance languages of western Eurasia have postnominal adjectives and gender systems, and the Chinese, Japanese and Korean languages of eastern Eurasia, which have prenominal adjectives and no agreement system, or gender).
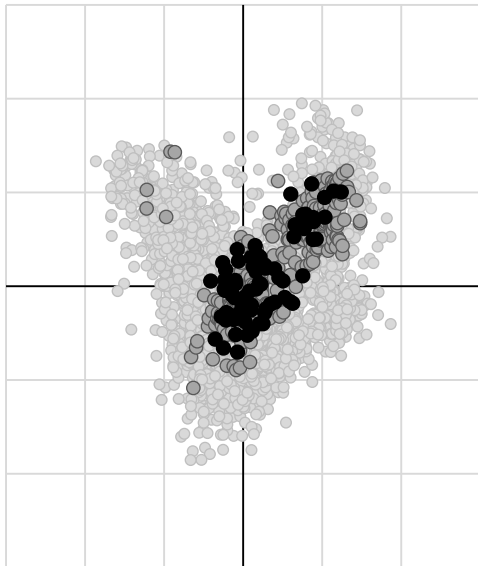
Nepali

(3)  सानो बिरालो रोयो
     sāno   birālo   ro-yo
     small   cat   cry-3SG.M.PAST
     'The small cat cried.'

(4)  केटी रोई
     keṭi   ro-i
     girl   cry-3SG.F.PAST
     'The girl cried.'

Tibetan

(5)  ཞི་མི་ཆུང་ཆུང་དུས་སོང་།
     zhimi   cungcung   ngü-song
     cat   small   cry-DISJUNCT
     'The small cat cried.'

(6)  བུ་མོ་དུས་སོང་།
     bumo   ngü-song
     girl   cry-DISJUNCT
     'The girl cried.'

Figure 2. *Dimensions 3 and 4*



This indicates that the four dimensions that are necessary to account for global variation can be reduced to two when considering only the languages of Nepal. That is to say, amongst the languages of Nepal we find that, for example, languages with case marking tend to lack elaborate verb agreement and applicatives (which follows from the correlation of Dimensions 1 and 2); similarly, the presence of plural marking or gender contrasts on nouns correlates with relative pronouns and modifiers that precede the noun that they modify (which follows from the correlation of Dimensions 3 and 4), but not with post-nominal modifiers in the NP, as seen in (3) – (6).

To better examine the variation among the languages of Nepal, we can plot them with a combination of Dimensions 1 and 2 along the *x*–axis, and a combination of Dimensions 3 and 4 along the *y*-axis. In the next section we see that language family divisions emerge naturally from such a plot, as well as some level of subgrouping which is reflected in the language typologies.

3. Variation in Nepalese languages

In Figure 3 we combine the two previous figures into one. The *x*-axis in this figure is defined as $(d_1 + d_2)/\sqrt{2}$, while the *y*-axis is defined as

$(d_3 + d_4)/\sqrt{2}$; as with Figures 1 and 2, the languages of Nepal (and immediate surrounds) are shown with black dots, additional languages of South Asia are shown in dark grey, and the other languages of the world, occupying a greater typological space, are marked with white dots. Since these combinations of axes were selected on the basis of their behaviour amongst the languages of Nepal, it is not surprising that these languages are more dispersed in Figure 3 than they were in either of Figure 1 or Figure 2. As with Figure 1, there is an extreme outlier in the right-hand part of Figure 3, namely Chhatthare Limbu, at approximately (+5, −1.5) in Figure 3. Given the dispersal of languages in Figure 3, we can visually identify three approximate clusters of Nepalese languages: one that is high on the *y*-axis (centred on approximately (−5, +6)), one that is low on the *y*-axis and low on the *x*-axis (approximately −7.5, 0), and one which is low on the *y*-axis but high on the *x*-axis at (0, −0.5), and which occupies the space between the second cluster just described, and Chhatthare Limbu.

Figure 3. *Combined axes*
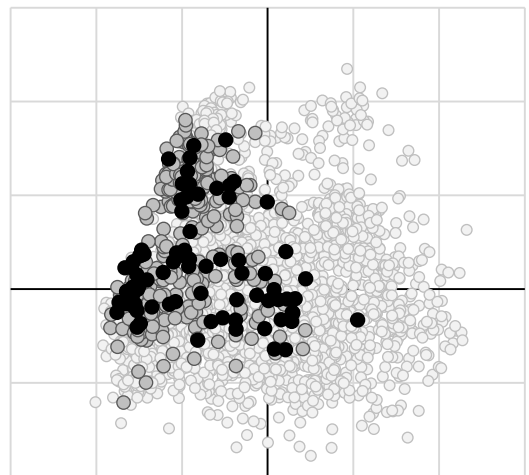


These three clusters correspond approximately to genealogical groups, as discussed below. In Figure 4 only the Nepalese languages from Figure 3 are displayed, with the Indo-European languages shown in black. (Figures 4–7 mark the axes in increments of two, so that the highest point indicated in Figure 4, representing Bhojpuri,

is at approximately (–2.5, +8).) It is clear that the first cluster identified above, high on the *y*-axis, occupies the typological space inhabited by the Indo-European languages of Nepal, characterised by high values on Dimensions 3 and 4 (corresponding to the presence of gender systems, plural marking on nouns, relative pronouns, prenominal modifiers, and suffixed verb agreement). We have seen examples of most of these features in (3) and (4); (7) illustrates the use of a relative pronoun in Rajbangshi (Poudel, 2006, p 73). Relative pronouns are almost completely unknown in languages towards the bottom of Figures 3 and 4, and are nearly ubiquitous amongst the Indic languages at the top of the figures.

Rajbangshi

(7)  dʒeira   ceŋra   dʒar
     who      boy     fever
     aːs-is-l-e…
     come-COMP-PT-3
     'the boy who had fever…'

The outlier, in the bottom left of Figure 4 near (–8, +1), is 'Hill Nepali', the variety of Nepali spoken by Tibeto-Burman speakers (lacking gender or verb agreement); it is effectively a relexified Bodic language. (Two other varieties of Nepali are included, at (–4, +5) and (–4.7, +4.9).) The Tharu languages are found around (–4.8, +5.6), and the four languages highest on the *x*-axis are split between Bhojpuri at (–2.5, +8) and a cluster of three southern Nepalese languages, Darai, Majhi and Maithili, around (–2.25, +5.5). The non-Indo-European languages closest to the Indo-European cluster are (from the top right) Santali (0, +4.5), the Munda language of south-eastern Nepal, and Kurukh (/Oraon), the Dravidian language of southern Nepal (–2.25, +4.9), close to the other southern Nepalese languages with typological profiles very typical of South Asia generally. We also find Tsum, the atypical Tibetan language from the Tsum valley in northern Nepal, at (–4.5, +3), closer to Indo-European because of its agreement system (Dhakal et al. 2016). These languages, similar typologically to the Indo-European languages without being related to them, represent long-term areal contact, showing the pan-South Asian

(indeed, pan-Central Eurasian) similarities between Indo-European, Dravidian, and Munda (see Donohue & Kalyan, 2024).

Figure 4. *The typological position of the Indo-European languages of Nepal*
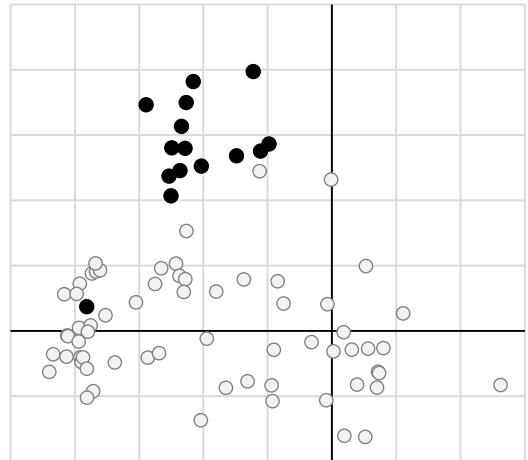


In Figure 5 we see the position of the TGTM (Tamang-Gurung-Thakali-Manang) languages, in grey, and the Tibetan languages, in black. These languages form the core of the cluster that is low on both the *x*- and the *y*-axes, presenting a very consistent typological profile, with a few exceptions. We have already discussed Tsum ((–4.5, +3), which is closer to the Indo-European languages than it is to the core of the TGTM-Tibetan cluster; between Tsum and the TGTM-Tibetan cluster is (Samagaun) Nubri (–5.5, +1.5), and the Tibetan language at furthest remove is Lower Nubri, from Prok (–2.75, +1.5); Lower Nubri, like Tsum, has a system of verbal agreement, a development of the conjunct/disjunct system that is robustly attested in Nubri varieties further up the valley. The less-distant language just below the zero line of the *y*-axis is Lhomi, the Tibetan language of the upper Arun valley (–5.5, –0.5). These languages have either reduced case systems, or else some level of agreement.

Amongst the TGTM languages the only significant outlier is Kuke, far to the right of the TGTM languages (–1.75, +1.5) and only minimally associated with them phylogenetically.

Kaike is the second most removed TGTM language at (−5.4, −0.7) (also acknowledged to be only distantly related to the core of the TGTM group) and Gurung at (−6.8, −1). Apart from these outliers, most of the TGTM languages are quite consistent in typology, and quite close to the Tibetan languages, though a little lower on the *x*-axis in Figure 5, reflecting their more consistent head-final typologies. Chantyal and Ghale are higher on the *y*-axis than the other members of this group (at around (−8, +1)), reflecting their longer and more intense histories of language contact. Example (8) shows a noun phrase from Chantyal (Noonan, 2005, p. 179). Here the indefinite *yewta* is borrowed from Nepali, as is its prenominal position.

Chantyal
(8)  yewta    cyãji    wadar-ri
     one      small    cave-LOC
     'in a small cave'

The Tibetan languages in this region of typological space are also languages that, while not being as divergent as the Tsum, Nubri and Lhomi languages, are heavily affected by contact with Indo-European languages: Gyalsumdo and Yohlmo, but also Shigatse from the Tibetan plateau (included as a language close to Nepal).

Figure 5. *The typological position of the TGTM and Tibetan languages of Nepal*
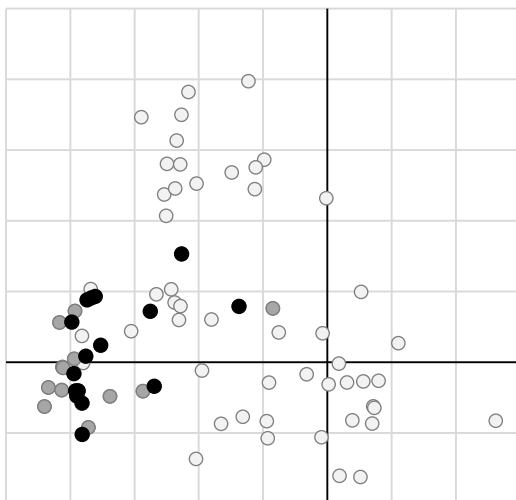


The languages not belonging to either of these two subgroups, but nonetheless occupying the same typological space, are (from the top of the bottom left cluster) Dura (−7.5, +2), Hill Nepali (as discussed above), and Baram (−7.5, 0). Close to this group is Dolakha Newari (−6, +0.9), between the TGTM-Tibetan cluster and the outlier Samagaun Nubri. The very typologically close group of languages between Samagaun Nubri and Tsum at (−4.5, +1.5) contains the other Newaric languages in the sample, and the Tibeto-Burman isolates Dhut Magar, Tinkar Lo and Dhimal, which despite being not closely related are typologically very similar to each other (see the discussion concerning Figure 7 below).

The more dispersed cluster to the right is shown in Figure 6. This cluster is spread out across a greater region of typological space than the other two clusters discussed here, though there is a clear core region occupied only by Kiranti languages. We have already mentioned Chhatthare Limbu as the South Asian outlier at the far bottom right of Figure 6 at (+5.25, −1.65). The other outliers are: to the left, Sunwar (−3.3, −1.75), Hayu (−2.6, −1.5) and Jero (−1.85, −2.15); at the top of the cluster, Puma (+1, +2) and Khaling (+2.2, +0.5); and just below the cluster at the bottom of the *y*-axis, Camling (+0.4, −3.2) and Dumi (+1, −3.2). The other eleven Kiranti languages in the study (see the appendix for a full list) are all similarly located in a tight cluster, with the typological dispersal similar to that in Tibetan or TGTM. Close to this cluster we find Kham (−0.1, +0.8), Chepang (−1.5, +0.8), Raute (−1.8, −0.5) and Raji (−1.9, −1.7). These languages are high on the *x*-axis as a result of being high on either or both of Dimensions 1 and 2. They display verb agreement, and have applicatives and modal marking on the verb, as in the following Chintang example (Schikowski et al., 2014, p.10). In (9) we see a verb with multiple agreement positions, and a malefactive applicative and directional marking, in addition to the verb root.

Chintang
(9)  Jamma
     all
     na-ca-i-ha-i-bir-i…

3>2-eat-2PL-AWAY.TR-2PL-APPL-2PL
'It might eat everything from you,…'

Figure 6. *The typological position of the Kiranti languages*
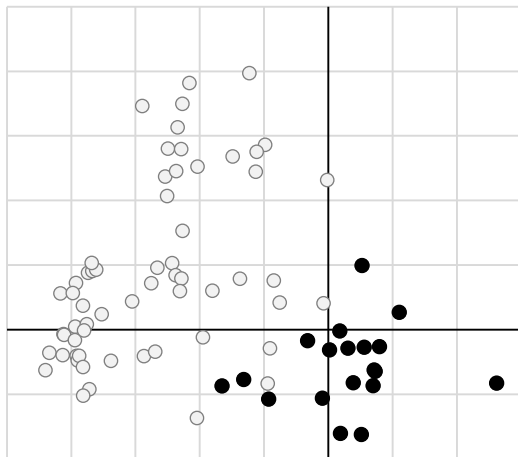


The remaining languages on the figures do not form a coherent group in typological space, though they are for the most part internal to the overall typological space. Three Newaric languages are shown with black circles; two Chepangic languages are shown with white circles; Magar is shown with a black triangle close to the Newaric languages, and Kusunda is shown with a grey triangle.

Figure 7. *Remaining Nepalese languages*
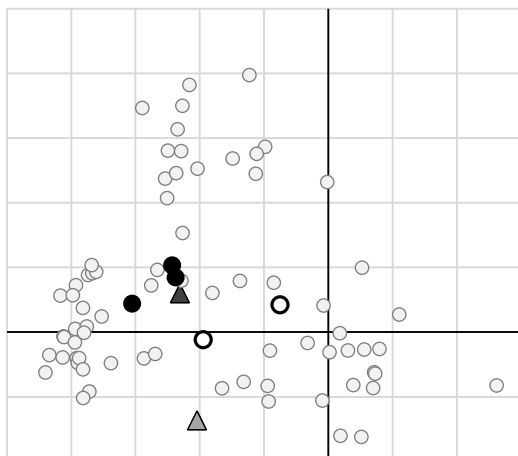


The Newaric languages are very close typologically to Magar, and also, as described earlier, Tinkar Lo and Dhimal. They occupy a middle ground between the Bodic group at the bottom left of Figure 7, and the Indo-European languages at the top. The Chepangic languages (Chepang and Bankariya) are removed from this group, further to the right due to their agreement systems, solidly within the group of Tibeto-Burman languages in the sample, but not particularly close to any other recognisable groups, though approaching the Kiranti cluster. The language isolate Kusunda is removed typologically from the other Nepalese languages surveyed, though not as much as Chhatthare Limbu is. Kusunda has a limited agreement system, but is found almost as far from the Indo-European languages as it is possible to be on the *y*-axis (Gautam, 2022), indicating a typological difference as great as is found between the Indic languages and the Kiranti languages.

4. The typological profile of Nepalese languages

We have seen that a typological analysis of the languages of Nepal, in a global context, reveals that, firstly, the language profile of Nepal is a restricted subset of global typological variation. This is to be expected in a country with only 1.5% of the world's languages. Indeed, for a such a small proportion of the languages, Nepal occupies a larger position in typological space than would be expected. The very low position on Dimension 1 (Figure 1) means that they profile of the languages of Nepal is dramatically skewed compared to the global profile, in that to a large extent the morphological profile of a language correlates with the (NP-internal) word order parameters. Once we allow for the typological profile of Nepalese languages, and adapt our display to better showcase this limited range of variation, we can see that the distribution of languages in typological space is predictable from knowledge of language family and linguistic subgroup. We have seen that the Indo-European languages are separated from the Tibeto-Burman languages (Figure 4), with exceptions being languages known to have been affected by linguistic contact. Amongst the Tibeto-Burman languages we can clearly identify a Bodish group

at the lower left of Figure 5, the Kiranti language groups at the lower right of Figure 6, and various unaffiliated languages in between these two areas. Again, exceptions are languages that are known, or suspected, to have undergone linguistic influence from neighbouring unrelated (or distantly related) languages. Despite the typological disruption caused by the many clear instances of language contact, the typological positions of the languages of Nepal correspond in broad strokes to their genealogical divisions.

Additionally, examining Figures 3 – 7 we can see that there is a region in typological space, at the top right of the figures, which is not populated by any languages of Nepal. These are the languages which simultaneously have elaborated verbal systems and also gender systems; examples include the Bantu languages of Africa, and many of the non-Pama-Nyungan languages of northern Australia.

Online materials

An interactive map that allows a visual appreciation of the position of the languages in typological space can be found at:

https://skalyan91.github.io/d3-language-maps/globe.html?data=ms-points-2023-noPC-Autotyp.

The map is best viewed with Google Chrome.

The coordinates of individual languages on dimensions 1 – 4 are available for download in the online materials for Kalyan and Donohue (2023), available at:
https://osf.io/u9qbe/

References

Dhakal, D. N., Donohue, M., Gautam, B. , & Liu, N. (2016). Diagnosing a contact history for Tsum. *Nepalese Linguistics, 31*, 14-20.

Donohue, M. (2019). A typological investigation of Nepalese languages. *Gipan, 4*(1), 30-37.

Donohue, M., & Kalyan, S. (2024). Defining a linguistic area from the bottom up: a pan-Eurasian perspective on areality and South Asia. [ Manuscript submitted for publication] *Journal of South Asian Languages and Linguistics*.

van Driem, G. (1987). *A grammar of Limbu*. Mouton de Gruyter.

Gautam, B. (2022/2079BS). *Kusunda bhasako byakarana* [A grammar of the Kusunda language]. Nepal Academy.

Kalyan, S., & Donohue, M. (2023). The dimensions of morphosyntactic variation: Whorf, Greenberg and Nichols were right. *Linguistic Typology at the Crossroads, 3*(2), 132-190. https://typologyatcrossroads.unibo.it/article/view/17482.

Noonan, M. (2005). Chantyal discourses. *Himalayan Linguistics Archive 2*, 1-254.

Poudel, T. (2006). *Rajbanshi grammar and interlinearized text*. LINCOM GMbH

Tumbahang, G. B. (2007). *A descriptive grammar of Chhatthare Limbu.* [Unpublished doctoral dissertation]. Tribhuvan University.

Appendix

List of languages of Nepal and surrounds examined in the article (name and iso code; where two or more lects share the same iso code they are differentiated by an added numeral).

Dravidian
Kurukh                          kxl

Isolate
Kusunda                         kgg

Indo-European; Indic
Awadhi                          awa
Hindi                           hin
Bhojpuri                        bho
Darai                           dry
Maithili                        mai
Majhi                           mjz
Chitoniya Tharu                 the
Danggaura Tharu                 thl
Rana Tharu                      thr
Saptariya Tharu                 thq
Kumauni                         kfy
Nepali                          nep
Nepali (hill)                   nep-2
Nepali (Bengal)                 nep-3
Nepali (Palpa)                  plp

Tibeto-Burman

'Kiranti'

| | |
|---|---|
| Bantawa | bap |
| Camling | rab |
| Puma | pum |
| Athpare | aph |
| Belhare | byw |
| Chintang | ctn |
| Yakkha | ybh |
| Yamphu | ybi |
| Bahing | bhj |
| Hayu | vay |
| Sunwar | suz |
| Jero | jee |
| Wambule | wme |
| Thulung | tdh |
| Dumi | dus |
| Khaling | klr |
| Koyi | kkt |
| Limbu | lif |
| Limbu (Chhathare) | lif-2 |

Tamang-Gurung-Thakali-Manange (+)

| | |
|---|---|
| Chantyal | chx |
| Gurung | gvr |
| Manange | nmm |
| Nar-Phu | npa |
| Seke | skj |
| Tamang | tge |
| Tamang (Eastern) | taj |
| Tamang (Western) | tdg |
| Thakali | ths |
| Ghale | ghe |
| Kuke | ght |
| Kaike | kzq |

Tibetan

| | |
|---|---|
| Dolpo | dre |
| Lhasa Tibetan | bod |
| Shigatse Tibetan | bod-2 |
| Kyirong Tibetan | bod-3 |
| Gyalsumdo | bod-4 |
| Jirel | jul |
| Yohlmo | scp |
| Lamjung Yohlmo | scp-2 |

| | |
|---|---|
| Lhomi | lhm |
| Mustang | loy |
| Nubri (upper) | kte |
| Nubri (lower) | kte-2 |
| Sherpa | xsr |
| Sherpa (Hile) | xsr-3 |
| Tokpe Gola Tibetan | ola |
| Tsum | ttz |

Chepangic

| | |
|---|---|
| Chepang | cdm |
| Bankariya | cdm-2 |

Newaric

| | |
|---|---|
| Kathmandu Newari | new |
| Dolakha Newari | new-2 |
| Pahari | new-3 |

Other Tibeto-Burman

| | |
|---|---|
| Thangmi | thf |
| Tinkar lo | bee |
| Raji | rji |
| Raute | rau |
| Baram | brd |
| Dhimal | dhi |
| Dura | drq |
| Kham | kgj |
| Dhut Magar | mgp |