UNRAVELING THE RELATIONSHIP AMONG THE KIRATI LANGUAGES

Tara Mani Rai

Central Department of Linguistics, T.U.
tara.rai@cdl.tu.np

This paper explores the relationships among the Kirati languages through lexical comparisons. The analysis, employing the Swadesh 100-word list, shows that Bantawa and Puma as well as Mugali and Phangduwali exhibit the highest lexical similarity, at 52%, while Yakkha and Koĩts-Sunuwar have the lowest similarity, at just 1%. In terms of phonetic similarity, Bantawa and Puma also show the greatest resemblance, with a similarity rate of up to 68%, whereas Mugali and Wambule show the least similarity, at 34%. These findings reveal that the lexical similarities and differences among the Kirati languages may reflect their geographical distribution and historicity.

Keywords: Koyee, Khaling, inclusive, exclusive, typological

1. Introduction

The Kirati languages belong to the Tibeto-Burman subfamily within the Sino-Tibetan language family. This classification encompasses languages spoken by communities such as the Rai, Limbu, Koĩts-Sunuwar, and Yakkha. Precisely, within the Rai-Kirati subgroup, there are 25 distinct languages, including Bantawa, Chamling, Khaling, Bahing/Bayung, Jero/Jerung, Wambule, Kulung, Thulung, Nachhiring, Dumi, Koyee, Sampang, Tilung, Puma, Dungmali, Lohorung, Yamphu, Mewahang, Sam, Athpare (Athpahariya), Chhintang, Chhiling, Belhare, Phangduwali, and Lungkhim (NSO, 2023). These languages are predominantly spoken in the eastern part of Nepal.

Hodgson (1857) is supposed to be the first study to gather linguistic data on the Kirati languages. It has compiled word lists for 17 distinct Kirati languages, including Bahingya [Bahing/Bayung], and Chourasya [*sic*. Unbule/ Wambule), Thulunggya [*sic.*Thulung], Khaling, Dumi, Rodong [*sic.*Chamling], Dungmali, Waling (Walahang), Rungchhenbung (clan name of the Bantawa) [latter three are Bantawa), Chhingtangya Sangpang [*sic.*Sampang], Nachhereng [*sic.*Nacchiring], Kulungya [*sic.* Kulung], Balali [*sic.* Mewahang], Lohoron [*sic.*Lohorung], Yakha [Yakkha], and Lambichhong [sic. Clanname of Yakkha]. Additionally, it has provided a grammatical description of the Bahing/Bayung language. Subsequently, Grierson (1909) and Shafer (1966) contributed to the classification of Kirati languages based on the work of Hodgson (1857, 1858).

A linguistic survey conducted in Nepal from 2009 to 2017 led by the Central Department of Linguistics at Tribhuvan University and the survey conducted by German Research Foundation (DFG) from 1981 to 1984 have gathered and analyzed data on the Kirati languages (Winter, 1986, 1987; Hanßon, 1991). Only after the 1970s, extensive studi es started on the Kirati languages like Thulung (Allen 1975), Bantawa (Rai 1985, Doornenbal 2009), Limbu (Weidert/Subba 1985; van Driem 1987, Dumi (van Driem 1993; Rai (2016), Athpahariya (Ebert 1997a, Neupane 2058 VS), Chamling (Ebert 1997b; Rai 2012), Yamphu (Rutgers 1998), Lahaussois (2002), Belhare (Bickel 2003), Wambule (Opgenort 2004a), Jero/Jerung (Opgenort 2005), Koĩts-Koĩts-Sunuwar (Rapacha 2005, Boörchers 2008), Kulung (Tolsma 2006), Chhatthare Limbu (Tumbahang 2007), Khaling (Jackques et al. 2012), Chhintang (Rai et al. VS 2067 [2011]), Puma (Sharma 2014), Koyee (Rai 2015), Chhintang (Poudyal 2015), Yakkha (Schackow 2015), Lungkhim (Rai et al. VS. 2078 [2020]).

Numerous publications on the Kirati language have been authored by native scholars, many of

which are written in Nepali. While substantial research exists on some Kirati languages, a few remain underexplored. Typological and structural studies of the Kirati languages include works by Michailovsky (1975), Genetti (1988, 1992), van Driem (1988), Ebert (1991), DeLancey (1992), Bickel et al. (2007), Jacques et al. (2012), van Driem (1993), Ebert (1993), Bickel (1999), Lahaussois (2003), Watters (2008), Allen (1972), Bickel (1997, 2001), Ebert (1999), Jacques and Lahaussois (2014), and Rai (2024).

The historical comparative studies, especially those focusing on lexical comparisons, have often yielded limited results. Of these, Opgenort (2005) identified fourteen lexical isoglosses across thirteen Kiranti languages, including Jero, Wambule, Bahing, Sunwar, Hayu, Khaling, Dumi, Thulung, Chamling, Bantawa, Kulung, Yamphu, and Limbu. In his subsequent study, Opgenort (2011) introduced twelve new lexical isoglosses, showing notable divergence from the 2005 findings. However, these additions were not strictly based on the Swadesh 100 wordlists. In contrast, this paper aims to explore the similarities and differences among the Kirati languages based on the Swadesh 100 wordlists.

2. Research methodology

The methods employed the Swadesh 100-word lists for lexical comparison. Initially, standardized Swadesh 100-word lists were gathered from various Kirati languages, including Jero/Jerung, Wambule, Thulung, Bahing/Bayung, Koĩts-Sunuwar, Sampang, Kulung, Nachhiring, Khaling, Koyee, Dumi, Tilung, Yakthung/Limbu, Yamphu, Lohorung, Mewahang, Lungkhim, Chmaling, Dungmali, Puma, Bantawa, Yakkha, Phangduwali, Mugali, Belhare, Aathpahare, Chhintang, and Chhiling.

These lists were compiled from the native speakers ensuring a diverse representation of gender, age, and literacy levels, and were validated by additional speakers from the same areas. The compiled wordlists were then entered into WordSurv (Wimbish, 1989), a tool used to assess the genetic relationships between languages. The data from WordSurv was exported as XML files and analyzed using Cog, a software that facilitates the comparison of languages through lexicostatistics and comparative linguistic techniques, automating much of the wordlist comparison process[1]. Secondary sources were also entertained during the study[2].

3. Findings and discussion

In this section, Swadesh 100-word lists from the Kirati languages are compared and analyzed using Cog, a recently developed software program for lexical and phonetic comparison.

In fact, the computer software Cog facilitates the researchers to compare and analyze wordlists from different language varieties using an elicitation approach. Using this program the researcher can quickly make sense of the data and then refine the wordlists and more settings, improving the comparison results and the understanding of the varieties at each step.

At first, the lexical similarity in percentage among the different Kirati langauge or the speech communities are presented and then shown phonetic similarities, namely Jero/Jerung, Wambule, Thulung, Bahing/Bayung, Koĩts-Sunuwar, Sampang, Kulung, Nachhiring, Khaling, Koyee, Dumi, Tilung, Yakthung/Limbu, Yamphu, Lohorung, Mewahang, Lungkhim, Chmaling, Dungmali, Puma, Bantawa, Yakkha, Phangduwali, Mugali, Belhare, Aathpahare, Chhintang, and Chhiling.

---

[1] The 60% cut-off point is commonly used as a benchmark for evaluating lexical similarity (Regmi 2013:63). However, this percentage should not be considered an inflexible standard. Speech varieties exhibiting lexical similarity below 60% are generally categorized as separate languages. In contrast, languages or dialects with a lexical similarity of 60% or more should be further assessed for intelligibility using the Recorded Text Test (RTT). Furthermore, the attitudes and perceptions of the speakers are also crucial factors to consider.

[2] The secondary data employed in this study is derived from (Kirat Rai Language and literature council (2062BS).

Table 1: *Lexical comparison among the Kirati groups*

As shown in Table 1, the highest degree of similarity, at 52%, is observed between Bantawa and Puma, as well as Mugali and Phangduwali. In contrast, the lowest degree of similarity, at only 1%, is found between Yakkha and Koĩts-Sunuwar.

Other notable similarities exceeding 40% include those between Nacchiring and Kulung (50%), Athapahare and Belhare (46%), Chhiling and Chhintang (45%), Bahing/Bayung and Koĩts-Sunuwar (44%), Belhare and Phangduwali (43%), Athapaher and Chhintang (43%), Nacchiring and Sampang (41%), Bantawa and Lungkhim (40%), Dungmali and Phangduwali (40%), Lohorung and Yamphu (40%), and Koyee and Dumi (40%). Conversely, the lowest similarity rates, ranging from 12% to 10%, are observed between languages such as Athapahare and Chhintang (3%) with Jero/Jerung, as well as among Chhintang, Chhiling, Mugaali, and Phangduwali, which each share 3% similarity with Phangduwali. Chamling, Yamphu, and

Lungkhim exhibit a 4% similarity with Wambule, while Athapaher, Belhare, and Lohorung share a 4% similarity with Koĩts-Sunuwar. Notably, these similarity ranges reflect the geographical distribution of the Rai-Kirati linguistic groups from west to east. It is evident that languages such as Jero/Jerong and Wambule from the Okhaldhung district, and Chhintang, Chhiling, Athapaher, Belhare, Mugaali, and Phangduwali from Dhankuta, do not share close cognates, likely due to their geographical distance.

Unless intelligibility testing is conducted using methods such as the Rapid Language Testing (RTT), lexical comparison alone with a 100 wordlists may not conclusively determine whether these are distinct languages or dialects. However, lexical comparison remains a crucial component in assessing whether a linguistic variety should be classified as a separate language or a dialect.

Table 2: *Phonetic similarity among the Kirati languages*

Unlike lexical similarity, phonetic similarity is not determined by slight variations in spelling or phonological patterns. Instead, phonetic similarity

reveals a higher degree of similarity in comparisons than lexical similarity does. As shown in Table 2, the Kirati languages exhibit a notable phonetic similarity.

Table 2 shows the phonetic similarity among Kirat languages based on 100 wordlists. The highest similarity, up to 68%, is observed between Bantawa and Puma, while the lowest is 34% between Muagali and Wambule. Other notable similarities of 60% and above include Athapahare and Belhare (67%), Muagali and Phangduwali (66%), Wambule and Jero/Jerumg (65%), Chhiling and Chhintang (65%), Bantawa and Dungmali (64%), Koyee and Dumi (64%), Nacchiring and Kulung (63%), Koĩts-Sunuwar and Bahing/Bayumg (62%), Phangduwali and Belhare (61%), Chamling and Puma (61%), Chhiling and Aathpahariya (61%), and Bantawa and Chhintang (60%).
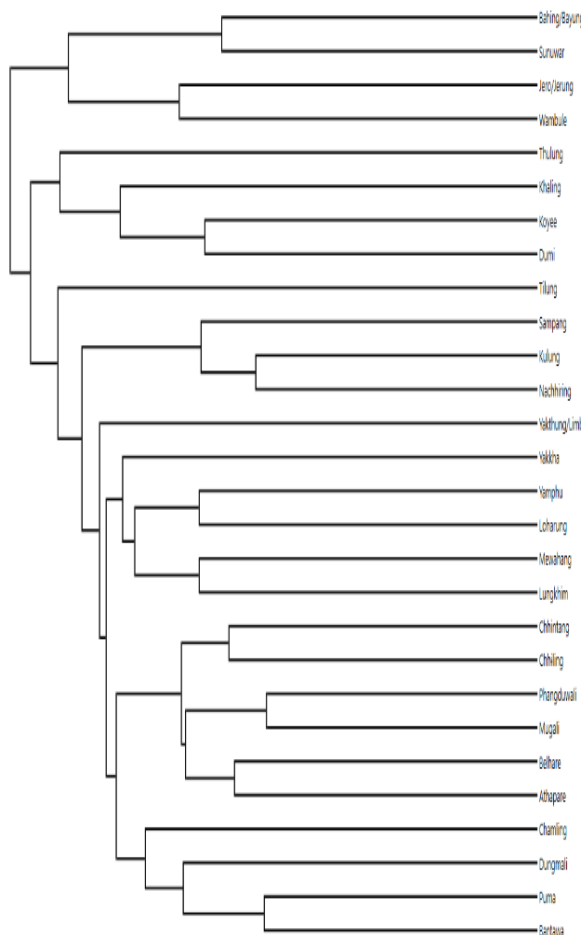
Figure 1:*Phonetic similarity among the Kirati lanauges*
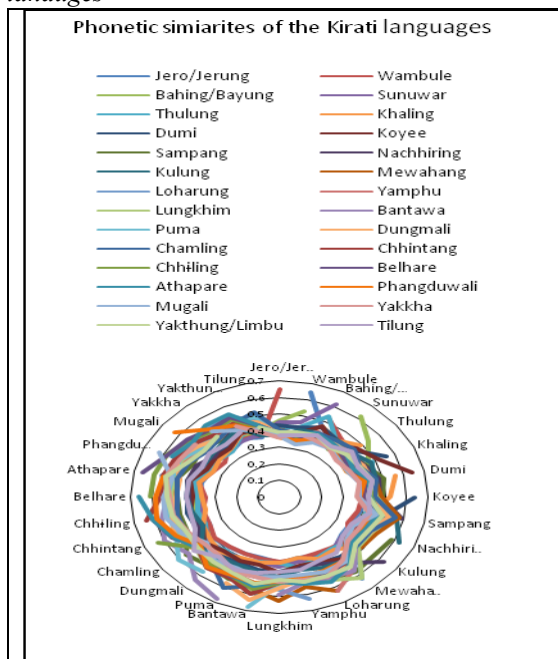
The lexicostatistical analysis of Kirati languages is represented through a hierarchical graph, which depicts the genetic relationships between various

language varieties based on their lexicostatistical similarities. To map these relationships, two primary methods are used: UPGMA (Unweighted Pair Group Method with Arithmetic Mean) and neighbor-joining. The UPGMA method constructs a rooted tree from a pairwise similarity matrix by iteratively merging the closest clusters into a single, higher-level cluster. Dendrogram 1 displays a hierarchical graph that visualizes this rooted tree structure.

Dendrogram 1:*Lexical similarity matrix*

Dendrogram 2: *Phonetic similarity matrix*
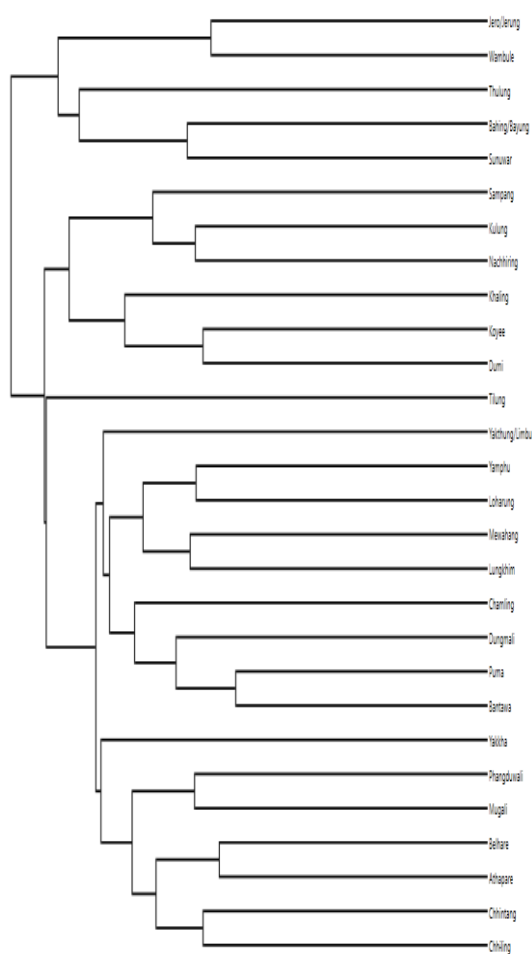
Dendrograms(1) and (2) illustrate the relationships among the Kirati languages, including Jero/Jerung, Wambule, Thulung, Bahing/Bayung, Koĩts-Sunuwar, Sampang, Kulung, Nachhiring, Khaling, Koyee, Dumi, Tilung, Yakthung/Limbu, Yamphu, Lohorung, Mewahang, Lungkhim, Chmaling, Dungmali, Puma, Bantawa, Yakkha, Phangduwali, Mugali, Belhare, Aathpahare, Chhintang, and Chhiling. The dendrogram reveals that many of these languages form pairs, such as Jero/Jerung with Wambule, Bahing/Bayung with Koĩts-Sunuwar, Nachhiring with Kulung, Koyee with Dumi, Yamphu with Lohorung, Mewahang with Lungkhim, Puma with Bantawa, Phangduwali with Mugali, Belhare with Aathpahare, and Chhintang with Chhiling. Conversely, while some languages share cognates with others, these relationships do not consistently form identifiable pairs.

4. Conclusion

The analysis of lexical similarity among the Kirati languages revealed that Bantawa and Puma as well as Mugali and Phangduwali exhibit the highest lexical similarity at 52%, whereas Yakkha and Koĩts-Sunuwar show the lowest similarity, at just 1%.

In terms of phonetic similarity, based on 100-word lists, Bantawa and Puma have the highest similarity, reaching 68%, while Mugali and Wambule show the lowest similarity at 34%. The dendrogram illustrating the relationships among Kirat languages, including Jero/Jerung, Wambule, Thulung, Bahing/Bayung, Koĩts-Sunuwar, Sampang, Kulung, Nachhiring, Khaling, Koyee, Dumi, Tilung, Yakthung/Limbu, Yamphu, Lohorung, Mewahang, Lungkhim, Chmaling, Dungmali, Puma, Bantawa, Yakkha, Phangduwali, Mugali, Belhare, Aathpahare, Chhintang, and Chhiling, demonstrates several notable pairings. For example, Jero/Jerung pairs with Wambule, Bahing/Bayung with Koĩts-Sunuwar, and Nachhiring with Kulung. However, while some languages share common cognates, these relationships do not always form clear and consistent pairs.

References

Allen, N. J. (1972). The vertical dimension in Thulung classification. *Journal of the Anthropological Society of Oxford, 3*, 81–94.

Allen, N. J. (1975). *Sketch of Thulung grammar*. Cornell University.

Bickel, B. (1996). *Aspect, mood, and time in Belharee*. Universität Zürich.

Bickel, B. (1997). Spatial operations in deixis, cognition, and culture: Where to orient oneself in Belhare. In J. Nuyts, & E. Pederson (Eds.), *Language and conceptualization* (pp. 46–83). Cambridge University Press.

Bickel, B. (1999). Nominalization and focus constructions in some Kiranti languages. In Y. P. Yadava & W. W. Glover (Eds.), *Topics in Nepalese linguistics* (pp. 271–296). Royal Nepal Academy.

Bickel, B. (2001). Deictic transposition and referential practice in Belhare. *Journal of Linguistic Anthropology, 10*, 224–247.

Bickel, B. (2003). Belhare. In G. Thurgood & R. J. LaPolla (Eds.), *The Sino-Tibetan languages* (pp. 546–570). Curzon Press

Bickel, B., Banjade, G., Gaenszle, M., Lieven, E., Paudyal, N. P., Rai, I. P., Rai, M., Rai, N. K., & Stoll, S. (2007). Free prefix ordering in Chintang. *Language, 83*(1), 43–73.

Borchers, D. (2008). *A grammar of Sunwar: Descriptive grammar, paradigms, texts, and glossary*. Brill.

DeLancey, S. (1992). Sunwar copulas. *Linguistics of the Tibeto-Burman Area, 15*(1), 31–38.

Doornenbal, M. A. (2009). *A grammar of Bantawa: Grammar, paradigm tables, glossary, and texts of a Rai language of Eastern Nepal*. LOT Publications.

Ebert, K. (1994). *The structure of Kirati languages: Comparative grammar and texts*. ASAS, Universität Zürich.

Ebert, K. (1997a). *A grammar of Athpare*. LINCOM EUROPA.

Ebert, K. (1997b). *Camling (Chamling)*. LINCOM EUROPA

Ebert, K. H. (1991). Inverse and pseudo-inverse prefixes in Kiranti languages. *Linguistics of the Tibeto-Burman Area, 14*(1), 73–79.

Ebert, K. H. (1993). Kiranti subordination in the South Asian areal context. In K. H. Ebert (Ed.), *Studies in clause linkage* (pp. 83–110). Arbeiten des Seminars für Allgemeine Sprachwissenschaft.

Ebert, K. H. (1999). The up-down dimension in Rai grammar and mythology. In B. Bickel & M. Gaenszle (Eds.), *Himalayan space: Cultural horizons and practices* (pp. 109–140). Völkerkundemuseum der Universität Zürich.

Genetti, C. (1988). Notes on the structure of the Sunwar transitive verb. *Linguistics of the Tibeto-Burman Area, 11*(2), 62–69.

Genetti, C. (1992). Segmental alternations in the Sunwari verb stem: A case for the feature [front]. *Linguistics: An Interdisciplinary Journal of Language Science, 30*(2), 317–358.

Grierson, G. (Ed.). (1909). *Linguistic survey of India: Tibeto-Burman family. General introduction, specimens of the Tibetan dialects, the Himalayan dialects, and the North Assam group* (Vol. 3.1).Motilal Banarsidass.

Hanßon, G. (1991). *The Rai of Eastern Nepal: Ethnic and linguistic grouping findings of the Linguistic Survey of Nepal* (W. Winter, Ed.). LSN and CNAS.

Hodgson, Brian Houghton (1858). On the Kiránti Tribe of the Central Himalaya. *Journal of the Asiatic Society of Bengal,* 27, 446–456.

Jacques, G., Lahaussois, A., Michailovsky, B., & Rai, D. B. (2012). An overview of Khaling verbal morphology. *Language and Linguistics, 13*(6),1095-1170.https://doi.org/10.21435/xxx.

Jacques, G., & Lahaussois, A. (2014). The auditory demonstrative in Khaling. *Studies in Language, 38*(2), 393–404.

Kirat Rai Langauge and Literature (2062 BS). *Kirati Bashaka swaesh sabdaharuko tulanatmak addhyan* [A comparative study of Swadesh lists in the Kirati languages]. Kirat Rai Yayokkha.

Lahaussois, A. (2002). *Aspects of the grammar of Thulung Rai: An endangered Himalayan language* (Unpublished doctoral dissertation). University of California.

Lahaussois, A. (2003). Nominalization and its various uses in Thulung. *Linguistics of the Tibeto-Burman Area, 26*(1), 33–57.

Michailovsky, B. (1975). Notes on Kirati verb- East Nepal-I: The Bahing verb-II: The origins of the tone in Khaling. *Linguistics of the Tibeto-Burman Area, 2*(2), 183-218.

Neupane, T. (2058 BS). *Athpahariya bhasako bhasik adhyayan* [A linguistic study of Athapariya] [Unpublished doctoral dissertation]. Tribhuvan University.

National Statistics Office/NSO. (2023). National report on caste/ethnicity, language & religion (Individual table): Table-2: Population by mother tongue and sex. https://censusnepal.cbs.gov.np/results/np/dow nloads/caste-ethnicity

Opgenort, J. R. (2004b). *A grammar of Wambule: Grammar, lexicon, texts, and cultural survey of a Kirati tribe of Eastern Nepal*. Koninklijke Brill.

Opgenort, J. R. (2005). *A grammar of Jero*. Brill.

Opgenort, J. R. (2011). A note on Tilung and its position within Kiranti. *Himalayan Linguistics, 10*(1), 253–271.

Paudyal, N. P. (2015). *Aspects of Chintang syntax* (Unpublished doctoral dissertation). Universität Zürich.

Rai, N. M. (2016). *A grammar of Dumi* [Unpublished doctoral dissertation]. Tribhuvan University.

Rai, N. K. (1985). *A descriptive study of Bantawa* [Unpublished doctoral dissertation]. University of Pune.

Rai, T. M. (2015). *A grammar of Koyee* [Unpublihed doctoral dissertation]. Tribhuvan University.

Rai, T. M. (2024). Clusivity in Chamling, Bantawa, Bayung, and Puma languages. *Tribhuvan University Journal, 39*(1), 81–96. https://doi.org/10.3126/tuj.v39i1.66676

Rai, T.M. (2078BS). *Lungkhim Bhasako ruprekha* [Sketch grammar of the Lungkhim langauge]. Report. Langauge Commission.

Rai, V. S. (2012). *A grammar of Chamling* [Unpublished doctoral dissertation]. Universität Bern.

Rai, N. K., Podyal, N., Robert, S., Gyanszle, M., Rai, M., Banjade, G., Rai, R., Rai, J. K., Rai, D., Rai, S., & Petigru, J. (2011). *Chhintang shabdokosh [Chhintang dictionary]*. Chintang Language Research Program, Tribhuvan University/University of Leipzig.

Regmi, A. (2013). *A sociolinguistic survey of Magar Kaike: A Tibeto-Burman language*. Report submitted to Linguistic Survey of Nepal (LinSuN)], Central Department of Linguistics, Tribhuvan University, Kathmandu, Nepal.

Rapacha, L. B. (2005). *A descriptive grammar of Kirant-Koits* [Unpublished doctoral dissertation]. Jawaharlal Nehru University.

Rutgers, R. (1998). *Yamphu: Grammar, texts, and lexicon*. CNWS Publications.

Schackow, D. (2015). *A grammar of Yakkha*. Language Science Press.

Shafer, R. (1966). *Introduction to Sino-Tibetan, Part I*. Otto Harrassowitz.

Sharma, N. P. (2014). *Morphosyntax of Puma, a Tibeto-Burman language of Nepal* (PhD thesis). SOAS, University of London. http://eprints.soas.ac.uk/18554

Tolsma, G. J. (2006). *A grammar of Kulung*. Brill.

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

van Driem, G. (1988). The verbal morphology of Dumi Rai simplicia. *Linguistics of the Tibeto-Burman Area, 11*(1), 134–207.

van Driem, G. (1993). *A grammar of Dumi.* Brill.

Widert, A., & Subba, B. (1985). *Concise Limbu grammar and dictionary*. Lobster Publication.

Watters, D. E. (2008). Nominalization in the Kiranti and Central Himalayish languages of Nepal. *Linguistics of the Tibeto-Burman Area, 31*(2), 1–44.

Winter, W. (1986). Bantawa rV- <?: An exercise in internal and comparative reconstruction. In D. Kastovsky & A. Szwedek (Eds.), *Linguistics across historical and geographical boundaries: In honour of Jacek Fisiak on the occasion of his fiftieth birthday* (Vol. 2, pp. 763–772). Mouton de Gruyter.

Wimbish, J. S. (1989). *Wordsurv: A program for analyzing language survey word lists*. Summer Institute of Linguistics.