

VERB ROOT IDENTIFICATION IN NEPALI: AN ALTERNATIVE PERSPECTIVE

BHIM NARAYAN REGMI

Central Department of Linguistics, T.U.

bhim.regmi@cbl.tu.edu.np

(Received: 09 Aug., 2025; revised: 09 Oct., 2025; accepted: 5 Nov., 2025; published: 26 Nov., 2025)

Infinitive suffix -nu removal technique from the citation form for basic and those with the high vowels in the root among the intransitive-transitive pairs with CaC ~ CaC and CuC ~ CoC structures for derived verb root identification among the contextual variants is practiced in Nepali. The other variants are related to the basic by vowel deletion and vowel lowering processes for basic and the derived roots respectively. This study presents an alternative perspective, i.e., the suffix -jo removal process to identify basic root and taking transitive root with low vowel as basic among the intransitive-transitive pair. The other variants are related to the basic variant by vowel insertion and vowel heightening processes for basic and derived intransitive root respectively. This study is based on the secondary data. Principle of majority and principle of precedence support this perspective, and the vowel insertion and vowel heightening processes are found more plausible.

Keywords: Verb root identification, non-permissible syllable structure, super heavy syllable, vowel insertion, vowel heightening

1. Introduction

Scholars working in Nepali language from early stages have proposed infinitive marker *-nu* (of citation form) removal technique to identify basic verb roots in Nepali (Ayton, 1820, p. 21; Turnbull, 1887, p. 17). Turner (1921, p. 101) has analyzed the infinitive markers *-nu*, *-ne*, *-na*, and *-na* of Nepali but he has said nothing about verb root identification. Later, he presented the entries on verbs with the infinitive marker *-nu* as citation form (Turner, 1931). Hem Raj Pandit in his grammar first published partly in 1969 BS follows the same procedure for verb root identification however he provides two different sets as examples together - first set which he calls roots includes the verb roots in the environment of the suffix *-nu* and the second set which he calls common forms include the roots in the environment of the suffix *-i* (See Pandit 2051BS, pp. 112, 120-25).

Scholars working in Nepali at the later stages such as Dahal (1974), Pokharel (2054BS), Adhikari (2055), Sharma (2071BS) also follow this basic root identification technique and present elaborate description of Nepali verbs.

Some Nepali verbs exhibit variations in their roots in the environment of the infinitive suffix *-nu* and the past, third person singular non-feminine non-honorific suffix *-jo*.

These variations include either a variation in a segment:

- /u/~o/ before suffix as in *d^hu-nu* 'to wash' and *d^ho-jo* 'washed',
- /u/~Ø before suffix as in *du^hu-nu* 'to milk', *du^h-jo* 'milked',
- /i/ or /a/~Ø before suffix as in *ughri/a-nu* 'to open', *ughr-jo* 'opened' and *kha^ltsi/a-nu* 'to pull' *kha^lts-jo* 'pulled',
- /i/ or /u/~Ø before suffix as in *kufi/u-nu*, *kuf-jo*,
- /a/~Ø before suffix as in *taula-nu* 'to weigh', *taul-jo* 'weighed',
- bisyllabic root (with metathesis and resyllabification) ~ monosyllabic root before suffix as in *pa^k.ga-nu* ~ *pa^k.ka^g-nu* 'to catch' and *pa^kg-jo* 'caught' etc.

Or a variation in the whole root:

-da~ga- as in *da-nu* 'go', *ga-jo* 'went'.

A group of derived Nepali verb roots with monosyllabic CVC structure varies in intransitive and transitive pairs. One of the members of the

pair is basic and the other is derived by a phonological process that changes the vowel height, such as *mar-* 'die' and *mar-* 'kill'. Variation in this class of roots is in a vowel as /ʌ/~a/, /i/~e/, and /u/~o/. Among these variants, the root with a high vowel is considered basic, and the root with a low vowel is considered derived in all the literature mentioned above. These roots are derived by the process vowel lowering.

However, Lohani (1999, pp. 49-54) has an alternative perspective regarding these roots. He argues that the roots with low vowels among these pairs are the basic, and the roots with high vowels are derived by the vowel heightening process. He has presented evidence from further derivation from the same root as well as from the derivation process in other word classes.

Whether a verb root has variations or not depends on its structure. Pokharel (2054BS, p. 9) has classified Nepali verb roots into the following five classes in terms of their structure: (1) consonant final - *gar-* 'do', *baral-* 'cause to roam', (2) diphthong final - *au-* 'come', *pau-* 'get', (3) monophthong final - *kʰa* 'eat', *li-* 'take', *de-* 'give', *dʰo-* 'wash', (4) multisyllabic with /h/ before final vowel - *kufi-* 'rot', *tsaha-* 'want', *kaha-* 'say', and (5) other multisyllabic - *pʰali-* 'spread', *orli-* 'climb down'.

Regmi (2025) has studied the variations in the citation forms of the verbs in dictionary entries. Though he has not included the variations in other forms of the verbs. He has discussed /u/~o/, /i/~e/, and /i/~u/ variations mentioned above in detail. He has classified the variations in the verb roots in the environment of the infinitive suffix, i.e., the citation forms, as syllable based variations and segment based variations. The syllable based variations include (a) variations of the verb roots containing front or back high vowel - /u/ ~ /i/ as in *kufu-nu* ~ *kufi-nu* 'to rot', (b) containing /ʌ/ or /i/ as in *utra-nu* ~ *utri-nu* 'to climb down, float', (c) presence or absence of /ʌ/ as in *taula-nu* ~ *taul-nu* 'to weigh', and (4) metathesis as in *pakḍa-nu* ~ *pakḍ-nu* 'to catch' which are explained in terms of syllable weight. He has presented a phonological explanation of these variations, but has not explicitly mentioned which one among the variants is basic.

The variations both in the basic and derived verb roots in Nepali mentioned above are observed based on three different types of evidences: (1) phonological environment, specially the initial segment of the suffixes, (2) frequency of all the forms of some selected verbs and other word classes derived from the same root, and (3) historical data on intransitive-transitive pairs of verbs derived by changing vowel height as well as the data from other classes of words. Based on these evidences, this study presents alternative explanation for the variations in Nepali verb roots.

2. Methodology

This study depends on secondary data. The sources include dictionaries, grammatical descriptions, and diachronic studies on Nepali. Regmi (2025) is used for the structural variations of Nepali verb roots in the environment of the infinitive suffix *-nu*, Sharma (2071BS), Poudel (2072BS) and a lemmatized version of Nepali National Corpus (NNCV2) are used for the variations in other environments, and Chalise (2071BS) is used for the historical data. These sources are the written texts in Devanagari except Regmi (2025). However, the data presented here as example are transcriptions of standard Nepali pronunciation in IPA. I converted the written data into spoken and transcribed myself as a native speaker of Nepali.

The study is developed on the idea that the basic among the variants of verb roots cannot be identified with the commonly practiced infinitive marker *-nu* removal process for basic verb roots, as well as with the vowel lowering process for derived verb roots. Instead, removing the suffix beginning with a vowel or an approximant for basic verb roots and the process of vowel heightening for derived roots.

In support of the idea or hypothesis, this article presents the selected data and arguments. Various types of reference works are consulted for both data and discussions.

The article presents two arguments in support of the hypothesis - (1) principle of majority - both the majority of a class of phonemes at the beginning of the inflectional and derivational

suffixes as well as the majority of occurrences of a class of phonemes at the beginning of the suffixes in the corpus, and (2) principle of precedence - the earlier evidences available in the historical records are basic.

3. The findings

Three sets of data are used in this article as mentioned earlier - inflections and derivations of the verb root ग् /gar-/ 'do' from a grammatical description and a dictionary, frequencies of lemmas घु /d^hunu/ 'to wash' and बस्नु /basnu/ 'to sit' from NNCV2, and the intransitive-transitive pair of verbs derived by changing vowel height from historical records. The data and the findings are presented in the following subsections.

3.1 Findings from the first set of data

All the inflections, derivations, and compound forms of the verb ग् /gar-/ 'do' are the first set of data taken from Sharma (2071BS, p. 286) and Poudel (2072BS, p. 156). This root with CVC structure shows no variation during any inflection or derivation process. Thus, it is selected for the identification of the suffixes - inflectional, derivational, and compounding. The initial phoneme of a suffix interacts with the final phoneme of the root. Thus, the initial phonemes of the suffixes mentioned above are identified and presented in Table 1 and 2. In some cases, phonemes other than the initial are also presented in the braces to make the presentation less abstract. These initial phonemes in the table are common elements of many suffixes¹. For example, *j* is the initial phoneme of the suffixes -*jo*, -*ja*u, -*ja*ũ.

Table 1: Initial phonemes of inflectional suffixes attached to the verb root gar-

Tense	Past	-j, -i, -e, -ẽ
	Non-past	-t ^h

¹ The term 'suffix' is loosely used in this article since the focus is on the initial phoneme of the element which immediately follows the verb root. This element may be suffix, compounding part, reduplicated part or any other elements which can be analyzed as an independent element such as auxiliary.

Aspect	Perfect	-e(k)
	Imperfect	-d
Mood	Habitual	-t ^h
	Inferential	-e(t ^h), -i(t ^h)
	Possibilitive	-l ² ,
	Imperative	-Λ
	Optative	-ũ, -Λũ, -e(s), -e, -o(s), -u(n)

Table 2: Initial phonemes of derivational suffixes attached to the verb root gar-

Noun	-ai, -au, -a(o), -o(t), -n(u), -n(Λ), -n(a)
Adjective	-u(wa), -e(k), -Λi(ja), -d, -n(e)
Adverb	-i, -e(Λ), -d(Λi), -n(Λ)
Verb	-au, -i

Compounding suffixes -*n(u)* and -*i* are attached to the root of the first member of the compound verb.

Initial phonemes of the inflectional, derivational, and compounding suffixes presented above still have repetitions. The number of initial phonemes which follow the verb root gar- without repetition is 18, including 6 consonants and 12 vowels (including 6 basic monophthongs, 2 nasalized and 4 diphthongs). These are /d/, /t^h/, /t^h/, /n/, /l/, /j/, /i/, /e/, /a/, /Λ/, /o/, /u/, /ẽ/, /ũ/, /ai/, /au/, /Λi/, /Λũ/. The suffixes beginning with a vowel and approximant, and the suffixes beginning with a consonant other than approximant, behave differently. Thus, these initial phonemes are grouped into two classes based on their phonological behaviour: (1) vowels and approximants (total 13), and (2) consonants other than approximants (total 5). Thus, the number of the first group of phonemes (vowels and approximants) is larger than the second group of phonemes (consonants other than approximants).

² The possibilitive marker in Nepali begins with /l/ that directly follows the root in the third person non-honorific and mid honorific forms. There are some other elements including agreement suffixes between root and the possibilitive marker in rest of the forms, such as gar-ũ-l, gar-ai-l, gar-au-l. Thus, there is a question whether it is a suffix or a cliticized form of any auxiliary.

3.2 Second set of data and findings

This set of data consists of the inflected and derived forms of two lemmas धुनु /d^hu-nu/ 'to wash' and बस्नु /bas-nu/ from the lemmatized version of Core Sample of Nepali National corpus (See, Hardie et al. 2011).

There are 41 instances of धुनु /d^hu-nu/ 'to wash' appeared as query result from the lemma search. While observing these results, three instances are of धुला /d^hula/ 'dust' which are not related to this verb. The remaining 38 instances of धुनु /d^hu-nu/ are related to 6 types - two with vowel as suffix initial phoneme and four with consonant as suffix initial phoneme. These are presented in Table 3a and 3b with their frequency of occurrences and the percentage.

Table 3a: धुनु /d^hu-nu/ with vowel initial suffix

	Word	IPA	Occurrences	Per cent
1.	घोए	/d ^h o-erΛ/	16	42.11%
2.	घोए	/d ^h o-e/	5	13.16%
	Total		21	55.26%

Table 3b: धुनु /d^hu-nu/ with consonant initial suffix

	Word	IPA	Occurrences	Per cent
1.	धुने	/d ^h u-ne/	9	23.68%
2.	धुनु	/d ^h u-nu/	5	13.16%
3.	धुन्छु	/d ^h u-(n)-ts ^h u/ ³	2	5.26%
4.	धुन्छ	/d ^h u-(n)-ts ^h Λ/	1	2.63%
	Total		17	44.74%

The verb presented in Table 3a and 3b has two variants of the root - d^ho- and d^hu-, which have vowel /e/ as suffix initial phoneme and consonant /n/ as suffix initial phoneme, respectively. Vowels at suffix initial position has 21 instances (55.26%) and consonants at suffix initial position has 17 (44.74%) instances out of total 38 instances.

बस्नु /bas-nu/ 'to sit' has 1531 instances appeared as query result from the lemma search. While observing these results, 13 instances are of a

particle बस् /bas/ 'no more' which are not related to this verb. The remaining 1518 instances of बस्नु /bas-nu/ are related to 37 types - 1 without suffix, 10 with vowel as suffix initial phoneme and 26 with consonant as suffix initial phoneme. These are presented in Table 4a, 4b and 4c with their frequency of occurrences and the percentage.

Table 4a: बस्नु /bas-nu/ without any suffix

	Word	Pronunciation	Occurrences	Per cent
1.	बस्	/bas/	6	0.40%

Table 4b: बस्नु /bas-nu/ with vowel or approximant initial suffix

	Word	Pronunciation	Occurrences	Per cent
1.	बसे	/bas-e/	399	26.28%
2.	बसेर	/bas-erΛ/	206	13.57%
3.	बसि	/bas-i/	90	5.93%
4.	बसी	/bas-i/	69	4.55%
5.	बस्यो	/bas-jo/	46	3.03%
6.	वसे	/bas-e/	22	1.45%
7.	बस्यौ	/bas-eu/	4	0.26%
8.	बसिन्	/bas-in/	2	0.13%
9.	बसुँ	/bas-ũ/	2	0.13%
10.	बसौँ	/bas-Λũ/	1	0.07%
	Total		841	55.40%

Table 4c: बस्नु /bas-nu/ with consonant (other than approximants) initial suffix

	Word	IPA	Occurrences	Percent
1.	बस्ने	/bas-ne/	231	15.22%
2.	बस्न	/bas-nΛ/	149	9.82%
3.	बस्नु	/bas-nu/	115	7.58%
4.	बस्दा	/bas-da/	41	2.70%
5.	बस्दै	/bas-dai/	30	1.98%
6.	बस्छन्	/bas-ts ^h Λn/	21	1.38%
7.	बस्छ	/bas-ts ^h Λ/	15	0.99%
8.	बस्थे	/bas-t ^h e/	14	0.92%
9.	बस्छु	/bas-ts ^h u/	9	0.59%
10.	बस्थेँ	/bas-t ^h ẽ/	7	0.46%
11.	बस्नै	/bas-nai/	6	0.40%
12.	बस्छे	/bas-ts ^h e/	5	0.33%
13.	बस्थ्यो	/bas-t ^h jo/	5	0.33%
14.	नबस्ने	/nΛ-bas-ne/	3	0.20%
15.	बस्छौँ	/bas-ts ^h Λũ/	3	0.20%
16.	बस्नोस्	/bas-nos/	3	0.20%
17.	बस्छ्यौ	/bas-ts ^h eu/	2	0.13%

³ Though /n/ between the root d^hu- and tense marker suffix -ts^h in serial numbers 3 and 4 in the Table 3b seems meaningless element in synchronic perspective, it can be explained as trace of habitual aspect marker suffix -d diachronically.

18.	बस्थी	/bʌs-tʰi/	2	0.13%
19.	बस्दैन्	/bʌs-dʌina/	2	0.13%
20.	बस्नुस्	/bʌs-nus/	2	0.13%
21.	बस्छौ	/bʌs-tʰʌu/	1	0.07%
22.	बस्थिन्	/bʌs-tʰin/	1	0.07%
23.	बस्थ्यौ	/bʌs-tʰeũ/	1	0.07%
24.	बस्दिन्	/bʌs-dinʌ/	1	0.07%
25.	बस्दैन्	/bʌs-dʌinan/	1	0.07%
26.	बस्लिस्	/bʌs-lis/	1	0.07%
	Total		671	44.20%

The verb presented in Table 4a, 4b, and 4c above has consonant final root which has no variation.

The verb form without suffix presented in Table 4a is the imperative non-honorific singular form of the verb, which is homophonic as well as homographic with the verb root.

The verb forms presented in Table 4b have vowels and approximant - /e/, /i/, /ʌ/, /ũ/ and /j/ as suffix initial phoneme and Table 4c have consonants - /n/, /d/, /tʰ/, /tʰ/ and /l/ as suffix initial phoneme. Among these, vowel at suffix initial position has 841 instances (55.40%) and a consonant at the suffix initial position has 671 (44.20%) instances out of a total 1518 instances.

Though there is a wide gap between the instances of occurrences of the verbs presented in Table 3a and 3b; and 4a, 4b, and 4c, there is similarity in the occurrences of vowel initial suffix and consonant initial suffix - roughly about 56% and 44% respectively.

The number of phonemes at the initial position of the suffixes in the first set of data shows the similarity with the second set of data, at least in the majority of the vowel and approximant phoneme initial suffixes over consonant (other than approximant) initial suffixes though these data are not directly comparable.

3.3 Findings from the third set of data

The third set of data belongs to the pairs of intransitive and transitive verb roots derived by changing vowel height. These data are taken from the inscriptions of Nepali from various times, even when the name of the language used to be Khas, Parbate, and Gorkha or Gorkhali.

The source of these data is Chalise (2071BS, pp. 155-68). He has discussed the causativization in Nepali based on the data from inscriptions. Among the processes of causativization, vowel height changing is under the scope of this article. He has also presented the analysis of the data; however, an anomaly can be seen there. At first, he discussed the causativization through vowel lowering, which derives transitive roots from intransitive roots by changing the vowels from /ʌ/ to /a/, from /u/ to /o/, and from /i/ to /e/. However, at another place (pp. 159-60), he has presented the same examples in a table format with the first available date of the roots in the inscriptions and explains the process as anticausativization through vowel heightening which derives intransitive roots from transitive roots by changing the vowels from /a/ to /ʌ/ and /o/ to /u/. He has also presented a case of vowel change from /e/ to /i/, but the place of example with vowel /i/ is left blank.

Two opposite explanations of the same data, without any clarification, show his confusion resulting from the available explanation in the literature and the data from the inscriptions, which are opposite to each other. Our concern here is what the data show instead of what his explanation is. Some of the examples are presented in (1) below.

- (1) a. Gloss Transitive Intransitive
 'Adhere' *lag* (1385BS) *lʌg* (1624BS)
 'Cut' *kaɽ* (1413BS) *kʌɽ* (1831BS)
 'Kill' *mar* (1497BS) *mar* (1497BS)
 'Wade' *tar* (1497BS) *tar* (1831BS)
 b. 'Put' *rakʰ* (1497BS) *rʌʃʌ* (1723BS)
 c. 'Leave' *tʰoɽ* (1534BS) *tʰuɽ* (1831BS)
 d. 'Join' *ɽoɽ* (1638BS) *ɽur* (1825BS)

The data in example (1) are systematically paired as transitive and intransitive verbs with /a/-/ʌ/, /o/-/u/, as well as /d/-/t/ and /d/-/r/ phonemes in the verb roots, respectively. The second member of each of the pairs is derived from the first member by anticausativization process of vowel heightening. Anticausatives are the intransitive verbs derived from transitive verbs.

In example (1), all the transitive verbs with /a/ or /o/ in the root in each of the pairs from the

different inscriptions belong to the earlier date, and the corresponding intransitive verbs with /ʌ/ or /u/ in the root belong to the later date. There is a difference of a minimum of 125 years to a maximum of 418 years and 260 years in average between the two members of the pair, except the one with meaning 'kill' in (1a), where both the verbs belong to the same inscription. Thus, the transitive roots are basic in terms of the principle of precedence.

There are also differences in consonants in addition to the vowels in the roots in (1b), (1c) and (1d) where the vowels /a/ and /o/ are followed by the consonants /kʰ/ and /g/ respectively where the vowels /ʌ/ and /u/ are followed by the consonants /hʌ/ and /tʃ/, /r/ respectively.

4. Discussion

As we have seen in the introduction section, the variations in the verb roots are of mainly two types - (1) variation in a part of the root that include - either presence or absence of a phoneme (u~Ø, i/ʌ~Ø, i/u~Ø, ʌ~Ø), either one or the other phoneme (u~o, ʌ~a), bisyllabic (with metathesis and resyllabification) ~ monosyllabic root, and (2) variation in the whole root (ɖa~ga).

Let's take type 2 first and settle it because it has nothing to do with phonological process. The two verb roots ɖa- and ga- have different historical sources. There are two roots in Sanskrit ja- 'go, visit, roam' and gam- 'go'. The Sanskrit loans जायार्क /jajabar/ 'nomad' and गमन /gaman/ 'visit' belong to these roots respectively. These two roots, with somewhat similar meanings, are supposed to be the variants of the same root in Nepali - the first one is used in non-past and the other in past tenses.

However, the tense based explanation is not so straightforward but erroneous too. Choice between these two forms is determined not by their meaning (tense in this case) but by their environment. The root ga- is used in the environment of the suffix beginning with a vowel or approximant consonant, and ɖa- is used in the environment of the suffix beginning with consonants other than approximants. Thus, there is ga- in perfective present tense form as in *gaeko*

tʰʌ, and there is ɖa- in imperfective past tense *ɖādai tʰijo*.

The role of environment is crucial with these verb roots, but the role is only on the choice between two verb roots, not on the change in the form, because these are not the variants of the same root but are different roots similar to the English verbs *go* and *went*. Solution to these variants is historical explanation.

The type 1 variations mentioned above need further discussion so continue in the following sections.

4.1 Defining verb root and environment

According to Katamba (1993, p. 41), root is 'the irreducible core of a word, with absolutely nothing else attached to it.'

Defining the environment of phonologically conditioned variations of any linguistic form is the departure for identifying the basic form among the variations. Phonological environment of any form is the preceding and following phonemes and boundaries, and structures. Nepali verb roots show variations at the end, not at the beginning. It indicates that the causes of variations are initial phonemes of the suffixes or other elements which are attach to the roots. These initial phonemes need to be observed in order to identify the environment of a verb root.

There is either presence or absence of a phoneme (u~Ø, i/ʌ~Ø, i/u~Ø, ʌ~Ø) among the type 1 variants of roots mentioned above. Following the definition of the root, if we take the variant in which a phoneme is present as part of the root, the other variant can be explained as the result of deletion. However, satisfactory explanation of this deletion process is missing in the literature.

The first and second sets of data show that the number of vowels and approximant consonants at suffix initial position is larger than the consonants other than approximants. This larger group of roots belong to the second member of the pairs u~Ø, i/ʌ~Ø, i/u~Ø, ʌ~Ø, i.e., no vowel at the root final position. It will be natural to follow the majority of the cases. In this case the variants of the verb roots with vowel or approximant consonant at suffix initial position as environment

should be taken as basic in terms of principle of majority. When this principle is followed, the roots without vowel at the end will be basic and the insertion process should be applied to get the other variants.

4.2 Vowel insertion

There are indications in the previous literature that explain why vowel insertion is necessary.

Dahal (1974, pp. 483-4), while classifying the verb stems, calls 'enunciating vowels' to those vowels which are present in some variants and absent in others. Enunciating means 'say or pronounce clearly', so these vowels are present with the forms which have difficulty in pronunciation, i.e., the roots with not permitted structures.

Pokharel (2039BS, pp. 134-5) also mentions that Nepali does not tolerate consonant clusters at word-final position, with examples from English loans as well as the same words from Sanskrit source in Hindi and Nepali differing in pronunciation. These examples include - (1) pronouncing /p^hʌstʌ/ and /bestʌ/ in Nepali for the English words *first* and *best* respectively, and (2) pronouncing the Sanskrit words इन्द्र /indrʌ/, मात्र /matrʌ/, गोत्र /gotrʌ/, अम्ल /ʌmlʌ/, and नम्र /nʌmrʌ/ in Nepali as in Sanskrit which are pronounced without the final /ʌ/ in Hindi.

Regmi (2025) has explained those not permitted syllable structures in the verb roots as super heavy syllables either with a CC cluster at coda position or with a diphthong as nucleus followed by C as coda or with /fi/ at coda position. Part of the examples (2-7) are taken from Regmi (2025), and the variants with the -jo suffix are added. These examples containing the roots with non-permissible syllable structures in Nepali as basic variants are placed at the right with the -jo suffix as the environment. The other variants precede this basic variant, which contain /i/, /ʌ/, or /u/ vowels inserted at the root final position to simplify the syllables. Example (7) exhibits metathesis and resyllabification in addition to vowel insertion.

u~Ø

- (2) a. 'milk' *duhu-nu ~ duhi-jo*

i/u~Ø

- (3) a. 'rot' *kuhi-nu ~ kuhu-nu ~ kuhi-jo*
 b. 'leak' *tsuhi-nu ~ tsuhu-nu ~ tsuhi-jo*
 c. 'abort' *tufi-nu ~ tuhu-nu ~ tufi-jo*
 e. 'gin' *ruhi-nu ~ ruhu-nu ~ ruhi-jo*

ʌ~Ø

- (4) a. 'say' *kʌhiʌ-nu ~ kʌhi-jo*
 b. 'tolerate' *sʌhiʌ-nu ~ sʌhi-jo*
 c. 'stay' *rʌhiʌ-nu ~ rʌhi-jo*
 c. 'flow' *bʌhiʌ-nu ~ bʌhi-jo*

Examples (2-4) have monosyllabic roots with /fi/ at coda position in the basic verb root, which is not permitted in Nepali syllable structure. Before the suffix -jo (or any other suffixes beginning with a vowel or approximant), these roots are resyllabified where the phoneme /fi/ shifts from the coda position of the first syllable to the onset position of the second, resulting in a permissible syllable structure; thus, there is no need of insertion.

However, before the suffix -nu (or any other suffix beginning with a consonant other than an approximant), either /fi/ deletion or vowel insertion is the option to get the syllable structure permissible. These examples show that Nepali chooses vowel insertion over the /fi/ deletion process. This process of insertion results into bisyllabic roots shifting /fi/ to the onset position of the second syllable from the coda position of the first syllable.

i/ʌ~Ø

- (5) a. 'climb down' *utri-nu ~ utra-nu ~ utr-jo*
 b. 'germinate' *umri-nu ~ umra-nu ~ umr-jo*
 c. 'roll down' *gultʌ-nu ~ gultʌ-nu ~ gult-jo*
 d. 'frighten' *tarsi-nu ~ tarsa-nu ~ tars-jo*
- (6) a. 'float' *tairi-nu ~ taira-nu ~ tair-jo*
 b. 'weigh' *tauli-nu ~ taula-nu ~ taul-jo*
 c. 'wear' *pʌiri-nu ~ pʌira-nu ~ pʌir-jo*
 d. 'swim' *pʌuɖi-nu ~ pʌuɖʌ-nu ~ pʌuɖ-jo*
 e. 'wake' *biũɖ^hi-nu ~ biũɖ^hʌ-nu ~ biũɖ^h-jo*
 f. 'be excited' *ɦʌusi-nu ~ ɦʌusa-nu ~ ɦʌus-jo*

Example (5) contains the basic variants in the environment of the suffix -jo with CC cluster at coda position presented at the right. This /j/ at

suffix initial position allows resyllabification that breaks the CC cluster to make the syllable structure permissible without insertion. Two other variants to the left are in the environment of the suffix *-nu*, where the vowels /i/ and /u/ are inserted. The vowel insertion changes the super-heavy monosyllabic roots into bisyllabic. This process breaks the CC cluster at the coda position into first C as coda of the first syllable and second C as onset of the second syllable. This change makes the syllable structure permissible.

Similarly, basic variants in the environment of the suffix *-jo* are presented at the right in example (6). These roots also have super-heavy monosyllables containing diphthongs as nucleus, followed by a consonant as coda, which is also not permitted in Nepali. The environment of the suffix *-jo* (or any vowel and approximant at suffix initial position) allows this super-heavy monosyllable to break into two syllables by shifting the consonant at the coda position of the monosyllable to the onset position of the second syllable. For the similar process of resyllabification, the other two variants need a vowel insertion process.

Λ + Metathesis and resyllabification ~ Ø

- (7) a. $\text{la} \sim \text{al}$
'climb up' *uk.la-nu* ~ *u.kal-nu* ~ *ukl-jo*
b. $\text{ra} \sim \text{ar}$
'climb down' *ut.ra-nu* ~ *u.tar-nu* ~ *utr-jo*
c. $\text{ʈa} \sim \text{aʈ}$
'attack' *ʈaʈp.ʈa-nu* ~ *ʈaʈa.paʈ-nu* ~ *ʈaʈa.p.ʈ-jo*
d. $\text{pa} \sim \text{ap}$
'capture' *ʱaʈd.pa-nu* ~ *ʱaʈa.dap-nu* ~ *ʱaʈd.p-jo*
e. $\text{ka} \sim \text{ak}$
'scold' *ʱaʈd.ka-nu* ~ *ʱaʈa.dak-nu* ~ *ʱaʈd.k-jo*
f. $\text{ga} \sim \text{ag}$
'catch' *paʈk.ga-nu* ~ *paʈa.kag-nu* ~ *paʈk.g-jo*

Example (7) contains the basic variants of verb roots with the CC cluster before the suffix *-jo* at the right. The roots at the left in the environment of the suffix *-nu* contain an inserted vowel /a/ together with metathesis and resyllabification. The roots at the left most position contain an inserted vowel /a/, which simplifies the structure by changing the super heavy syllable into two permissible syllables.

This process is plausible but there is further change in the variants presented in the middle.

The inserted vowel and the preceding consonant change their places with each other in these examples which results into resyllabification. This process applied to already permissible syllable structures has no answer at the moment, which may need further investigation.

The verb root identification in the environment of the suffix with a vowel or approximant not only follows the majority principle but also supports logical explanation.

4.3 Vowel heightening

Findings of the third set of data presented above set the ground for considering the transitive root basic. Earlier occurrences of the roots with lower vowels as the nucleus is one of the strong evidence to consider these roots as basic variants, as presented in example (1).

Intransitive derivation by heightening the vowel of the transitive root is also supported by the following examples. Vowel heightening as a derivation process considers the transitive verb form basic and results into both intransitive and causative as shown with the single headed arrow in example (8).

(8) Gloss	Intransitive	Transitive	Causative
'Adhere'	<i>lag</i> ←	<i>lag</i> →	<i>lag-a</i>
'Cut'	<i>kaʈ</i> ←	<i>kaʈ</i> →	<i>kaʈ-a</i>
'Kill'	<i>mar</i> ←	<i>mar</i> →	<i>mar-a</i>
'Wade'	<i>tar</i> ←	<i>tar</i> →	<i>tar-a</i>
'Put'	<i>raʱa</i> ←	<i>raʱh</i> →	<i>raʱh-a</i>
'Leave'	<i>ʈaʱu</i> ←	<i>ʈaʱo</i> →	<i>ʈaʱu-a</i>
'Join'	<i>ʈur</i> ←	<i>ʈo</i> →	<i>ʈur-a</i>

The differentiating factor between the intransitive and the causative is that the causative takes the suffix *-a* in addition to the vowel heightening process.

When the intransitive root is considered basic and vowel lowering the derivation process, it results into transitive. However, causative derivation from this transitive needs vowel heightening in addition to suffixation, as in example (9). This is the widely available explanation in Nepali literature.

(9) Gloss	Intransitive	Transitive	Causative
'Adhere'	<i>lag</i> →	<i>lag</i> →	<i>lag-a</i>
'Cut'	<i>kaʈ</i> →	<i>kaʈ</i> →	<i>kaʈ-a</i>

'Kill'	<i>maɾ</i>	→	<i>maɾ</i>	→	<i>maɾ-a</i>
'Wade'	<i>taɾ</i>	→	<i>taɾ</i>	→	<i>taɾ-a</i>
'Put'	<i>raɦa</i>	→	<i>raɦ^h</i>	→	<i>raɦ^h-a</i>
'Leave'	<i>ʈ^huɽ</i>	→	<i>ʈ^hoɽ</i>	→	<i>ʈ^huɽ-a</i>
'Join'	<i>ɽur</i>	→	<i>ɽoɽ</i>	→	<i>ɽur-a</i>

Example (9) has the same data as presented in (8), but the derivation process is different. As the arrows indicate, intransitive roots are basic, and the derivation process is vowel lowering to derive transitive, but the process is vowel heightening to derive causative from transitive.

Considering transitive roots with a lower vowel as the nucleus, the basic and derivation process of vowel heightening is simple because it applies to a transitive root to form both the intransitive and causative. But considering the intransitive root basic and vowel lowering as derivation process is complex because we need a vowel heightening process for further derivation from the transitive to the causative. This explanation is implausible which lacks the directionality of a phonological process. Simple over complex, and one process over two processes is a more effective description.

The process of vowel heightening is also supported by the feminine noun derivation from a non-feminine surname or caste/ethnicity in the Nepali community.

(10)	Gloss	Non-feminine	Feminine
	'Thapa'	<i>thapa</i>	→ <i>thapi-ni</i>
	'Karki'	<i>karki</i>	→ <i>karki-ni</i>
	'Rai'	<i>rai</i>	→ <i>rai-ni</i>
	'Kami'	<i>kami</i>	→ <i>kami-ni</i>
	'Sarki'	<i>sarki</i>	→ <i>sarki-ni</i>
	'Tharu'	<i>t^haru</i>	→ <i>t^haru-ni</i>
	'Lama'	<i>Lama</i>	→ <i>lami-ni</i>

In example (10), feminine nouns are derived by heightening of the vowel in the first syllable of the basic non-feminine nouns in addition to the suffix *-ni*.

Lohani (1999, pp. 49-54) has discussed the root identification problems in *CaC ~ CaC* pairs, which can be extended to *CiC ~ CeC* and *CuC ~ CoC* pairs too. He has similar arguments as presented in examples (8) and (9) above. Furthermore, he has presented the examples of the

vowel heightening process in deriving other word classes as presented in (11).

(11)	<i>kaɦ</i> 'work'	<i>kaɦai</i> 'earning'
	<i>g^ham</i> 'sun rays'	<i>g^hamaɦlo</i> 'full of sun rays'
	<i>laɽ</i> 'shame'	<i>laɽalu</i> 'shameable'
	<i>laɽ</i> 'shame'	<i>laɽau</i> 'feel shame' (p. 51)

Examples in (11) show a similar process as in examples in (10).

Sanyal, Sharma, and Prasad (2021) have also observed a similar pattern of verb root allomorphy and explain the process as vowel reduction based on the sonority hierarchy. They have taken the transitive root as $\sqrt{R1}$ and both the intransitive and indirect causative roots as $\sqrt{R2}$. Their example from Hindi is presented here as (12).

(12)	Intransitive	Transitive	Indirect Causative
	<i>kəɽ-na:</i> 'to be cut'	← <i>kaɽ-na:</i> 'to cut'	→ <i>kaɽ-wa:-na:</i> 'to cause one to cut'
	<i>juɽ-na:</i> 'to be joined'	← <i>joɽ-na:</i> 'to join'	→ <i>juɽ-wa:-na:</i> 'to cause to be joined'
	<i>luɽ-na:</i> 'to be robbed'	← <i>lu:ɽ-na:</i> 'to rob'	→ <i>luɽ-wa:-na:</i> 'to cause to be robbed'

(Sanyal, Sharma, & Prasad, 2021, p. 395)

These evidences from historical evidence follow the principle of precedence, and the vowel heightening process available in other word classes also supports the analysis that derives intransitive from basic transitive roots by the vowel heightening process. A similar explanation for Hindi data also supports this analysis.

The process can be called 'lowering', however, the term 'first formant' should be chosen instead of 'vowel' because it is the first formant, not the vowel that undergoes the process.

4.4 The basic verb roots

Following the findings and discussion above, Nepali verb roots need to be reclassified based on the environment of the suffix *-jo*. For this purpose, we can take the classification by Pokharel (2054BS, p. 9) as a reference and make necessary changes as follows:

(1) Consonant final - *gar-* 'do', *baral-* 'cause to roam' can be renamed as single consonant final roots.

(2) Diphthong final - *au-* 'come', *pau-* 'get' can be renamed as vowel final and grouped with class (3) monophthong final - *k'a* 'eat', *li-* 'take', *de-* 'give', *d^ho-* 'wash'. These can be subgrouped as those insert /ũ/ and those insert /n/ before the suffix -*ts^h*.

(4) Multisyllabic with /fi/ before final vowel - *kufi-* 'rot', *tsafiA-* 'want', *kafiA-* 'say', can be renamed as /fi/ final roots.

(5) Other multisyllabic - *p^haili-* 'spread', *orli-* 'climb down' need to be renamed as roots with super heavy syllables, which can be subdivided into those with diphthong followed by a consonant and those with consonant cluster final.

This reclassification is as follows:

(1) Single consonant final roots with monophthong.

(2) Vowel final roots - (a) roots which insert /ũ/ before the suffix -*ts^h*, and (b) roots which insert /n/ before the suffix -*ts^h*.

(3) /fi/ final roots.

(4) Roots with super heavy syllable - (a) single consonant final roots preceded by a diphthong, and (b) consonant cluster final roots.

After reclassifying the verb roots, the variations in a phoneme within the roots are addressed as follows.

The root with low vowel /o/ is considered basic among the /u~/o/ variants before the suffix, as in *d^hu-nu* 'to wash' and *d^ho-jo* 'washed', and the root with high vowel /u/ is derived by the vowel heightening process. These roots belong to class 2a.

The root without vowel at the end is considered basic among the /u~/∅ and /i/ or /u~/∅ variants before suffix as in *dufu-nu* 'to milk', *duh-jo* 'milked', and as in *kufi/u-nu* 'to rot' *kufi-jo* 'rotted' respectively. The variant with /i/ or /u/ is derived by the vowel insertion process. These roots belong to class 3. These classes of roots can also insert any other vowel instead of /u/.

The root without vowel at the end is considered basic among the variants /i/ or /ʌ~/∅ and /ʌ~/∅

before suffix as in *ughri/ʌ-nu* 'to open', *ughr-jo* 'opened', *khāṭsi/ʌ-nu* 'to pull' *khāṭs-jo* 'pulled', and *taula-nu* 'to weigh', *taul-jo* 'weighed'. The variant with the vowels /i/ or /ʌ/ is derived by the vowel insertion process. These roots belong to class 4.

The monosyllabic root is considered basic, and the bisyllabic root is derived by vowel insertion followed by metathesis among the bisyllabic roots (with metathesis and resyllabification) ~ monosyllabic root before suffix as in *pak.qa-nu* ~ *pʌ.kʌq-nu* 'to catch' and *pakq-jo* 'caught' etc. This class of roots also belongs to class 4.

All the variants except those that belong to type 2 roots consider the roots without vowel at the end as basic and apply the vowel insertion process for derivation. The insertion is purposeful as the process changes the syllable structure to make it permissible.

5. Conclusion

Scholars working at the early stages in Nepali grammar proposed the infinitive suffix -*nu* removal technique for basic verb root identification, and those working at the later stages have followed it and presented elaborate descriptions of the various forms of the verbs, variations in the verb roots, and the phonological processes to explain the variations. On a phonologically derived set of intransitive-transitive pairs with CʌC~CaC structure, the intransitive root is considered basic, and the transitive root is derived by vowel lowering process.

This study presents an alternative perspective to verb root identification, i.e., removal of the third person non-honorific non-feminine singular past tense suffix -*jo* as representative of a group of suffixes beginning with a vowel or approximant.

Two sets of data on basic verb roots and one set of data on derived verb roots have been presented in support of this alternative perspective. The first set of data includes the number of phonemes at the beginning of the inflectional, derivational, and compounding suffixes attached to the verb root. The second set of data includes the frequency of various forms of words related to the same

lemma. These two sets of data show that the number of vowel and approximant phonemes is larger than the remaining phonemes (consonants other than approximants) at the initial position of the suffixes attached to the verb root. These findings from the data comply with the principle of majority, which is in favour of the *-jo* removal technique for basic verb root identification.

The third set of data includes the inscriptional records of intransitive-transitive pairs of derived verb roots of C₁C ~ CaC and CuC ~ CoC syllable structure with the dates of their first occurrences. This set of data shows that the transitive roots with CaC and CoC structure first occurred in the inscriptions earlier than their intransitive counterparts with C₁C and CuC structure. The principle of precedence supported the idea that considers the transitive root basic and the intransitive root derived by the vowel heightening process.

This alternative perspective on both the basic and derived roots is also supported by other data in which similar processes are applied, as well as by the literature on similar data in Hindi. This alternative perspective offers more regular and efficient explanations of verb root variations.

This alternative perspective on the root identification technique requires reclassification of the basic verb roots which results into the following four classes: (1) single consonant final roots, (2) vowel final roots, (3) /h/ final roots, and (4) roots with heavy syllables. Class 2 roots are further classified as /ũ/ inserting and /n/ inserting roots before suffix *-ʰ*, and class 4 roots into single consonant preceded by diphthong and CC cluster at final.

The roots identified by the suffix *-jo* removal technique undergo only two processes - (1) class 2 basic verb roots and derived verb roots from the CaC transitive roots undergo vowel heightening process while deriving C₁C intransitive roots, and (2) other basic verb roots undergo vowel insertion process in the environment of the suffixes beginning with consonant other than

approximants in order to make the syllable structure permissible.

Results of this alternative technique of root identification comply with Katamba's definition of root as 'the irreducible core of a word, with absolutely nothing else attached to it', whereas the *-nu* removal technique leaves the additional materials which are used for making the structure fit into the phonological system of the Nepali language as if these are the parts of a root.

The findings from this study suggest the need for revision in citation forms of Nepali verbs in the dictionary entry.

Finally, the verb root is an analytical unit not an independent form so it is an underlying abstract unit which can be ready to be an independent form only after some affixes - either overt or covert - attached to it. In course of getting an independent form, the roots with non-permissible structures undergo phonological processes whereas the roots with permissible structure remain unchanged. This article explored such phonological processes which help in identifying the verb roots.

Acknowledgements

I would like to express my thanks to the participants who provided the valuable comments and suggestions when it was presented as paper in the same title in the joint conferences 5th Himalayan Languages Symposium and 20th Annual Conference of Linguistic Society of Nepal, in Kathmandu, 13-15 September 1999 and presented in the title 'Variations in Nepali verb root and question on citation form' in the joint conferences 36th South Asian Languages Analysis (SALA) Roundtable and 43rd Annual Conference of Linguistic Society of Nepal in Kathmandu, 26-29 November 2022. I would also like to extend my thank to the anonymous reviewers who read this article thoroughly and provided valuable comments and suggestions as well as the editorial team for their effort on through processing and editing.

References

- Adhikari, H. R. (2055BS). *Samasāmayika Nepālī vyākaraṇa* [Contemporary Nepali grammar] (2nd ed.). Vidyarthi Pustak Bhandar.
- Ayton, J. A. (1820). *A grammar of the Nepalese language*. Philip Pareira, Hindoostanee Press.
- Chalise, B. (2071BS). *Abhilekhīya Nepālī kriyā* [Inscriptional Nepali verbs] (2nd ed.) Unique Educational Publishers.
- Dahal, B. M. (1974). *A description of Nepali: literary and colloquial* [Unpublished doctoral dissertation]. University of Pune.
- Hardie, A., Lohani, R. & Yadava, Y. (2011). Extending corpus annotation of Nepali: advances in tokenisation and lemmatisation. *Himalayan Linguistics*, 10(1) [Special Issue in Memory of Michael Noonan and David Watters], 151–165, <https://doi.org/10.5070/H910123572> <https://cqpweb.lancs.ac.uk/nncv2>
- Katamba, F. (1993). *Morphology*. Macmillan.
- Lohani, R. R. (1999). *Complex predicates in Nepali* [Unpublished master's thesis]. Tribhuvan University.
- Pandit, H. (2051BS). *Chandrikā: Gorakhābhāṣā-vyākaraṇa* [A grammar of Gorkha language] (3rd ed.). Sajha Prakashan.
- Pokharel, B. (2039BS). *Nepālīkā kehi dhvanitāttvika Behorāharu* [Some phonological characteristics of Nepali]. In B. Pokharel (Ed.). *Pachchisa varṣakā bhāṣika charchā* [Discussions on language during 25 years] (pp. 123-42). Royal Nepal Academy.
- Pokharel, M. P. (2054BS). *Nepālī vākya vyākaraṇa* [Nepali Syntax]. Royal Nepal Academy.
- Poudel, M. P. (2072BS). *Nepālī kriyāharuko kośa (prayogātmaka)* [A dictionary of Nepali verbs (practical)]. Vidyarthi Prakashan.
- Regmi, B. N. (2025). Variation in Nepali verb roots: A phonological perspective. *Gipan*, 7(1), 121–135. <https://doi.org/10.3126/gipan.v7i1.84240>
- Sanyal, P., Sharma, V. & Prasad, A. (2021) Hindi root allomorphy: insights from phonological and morphosyntactic theory. In G. Sharma & J. J. Lowe (Eds.). *Trends in South Asian Linguistics* (pp. 381-414). De Gruyter Mouton. <https://doi.org/10.1515/9783110753066-015>
- Sharma, M. R. (2071BS). *Prajñā Nepālī sandarbha vyākaraṇa* [Prajñā Nepali reference grammar]. Nepal Academy.
- Turnbull, A. (1887). *A Nepali grammar and English-Nepali & Nepali-English vocabulary (about 4000 words): designed for the use of missionaries, tea-planters and military officers*. N. L. Roy.
- Turner, R. L. (1921). The infinitive in Nepālī. *Philologica: Journal of Comparative Philology*, 1, 101-116.
- Turner, R. L., Sir. (1931). *A comparative and etymological dictionary of the Nepali language*. Kegan Paul.