

A Comprehensive Study of Lepcha Monosyllabic Structure and Components

Ugen Lepcha

Assistant Professor, Department of Lepcha, Khangchendzonga State University, Sikkim

Abstract

This article focuses on the structure, classification, and linguistic functions of monosyllabic units of the Lepcha language. Special attention is given to the issue of whether a monosyllabic form is regarded as an independent word, meaningful morphological unit, grammatical or phonological marker, or simply a possible sound representation within the Lepcha writing system. The study deals with a comprehensive inventory of 11,548 monosyllabic forms identified in the Lepcha script. Their analysis follows the basis of meaning, grammar, phonetics, morphology, and interaction with diacritic marks. The forms were classified into two categories: 8,678 lexical and 2,870 non-lexical units. Lexical units include those that perform independent word functions, add meaning as parts of larger words, and fulfill grammatical and marking purposes. Non-lexical units, in turn, do not necessarily have an independent lexical meaning, but they are important as they stand for possible sound combinations or derive from specific letter-diacritic combinations. The article argues that monosyllabic forms are not understood simply as isolated words. They form integral parts of the Lepcha vocabulary, morphology, phonology, orthography, and language documentation. The difference between lexical and non-lexical forms is particularly relevant to dictionary-making and grammar, spelling, and language teaching studies, as well as to electronic, linguistic, and language preservation resources. The inventory, therefore, lays the groundwork for further studies into the Lepcha sound system, word formation, orthographic conventions, grammatical restrictions, and lexicographic documentation.

Keywords: Lepcha language, monosyllabic structure, diacritics, lexical units, phonetic representation, morphology, orthography, language preservation

Introduction

Language is not merely a means of communication, but it is also an important carrier of cultural identity, collective memory, knowledge, and tradition. For languages with comparatively limited documentation, systematic linguistic study is particularly significant. The Lepcha language spoken mainly in Sikkim and Darjeeling possesses a distinctive writing system and a linguistic structure that offers opportunities for the study of sounds, words, morphology, and orthography. The documentation and analysis of its linguistic units are therefore important both for linguistic scholarship and for the continued development and preservation of the language.

Published: 30 September 2026

DOI: <https://doi.org/10.70558/IJMRS.2026.v2.i3.301212>

Copyright © 2026 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

One particularly important feature of Lepcha is the presence and extensive use of monosyllabic forms. A monosyllabic form consists of a single syllable, but a single syllable does not necessarily perform only one linguistic function. Depending on its use, a monosyllabic form may occur as an independent word, as part of a larger word, as a prefix or suffix, as a grammatical or marking element, or as a representation of a possible sound. Consequently, the study of monosyllabic units requires attention to more than their physical written form.

The present study undertakes a study of Lepcha monosyllabic forms. A total of 11,548 forms were identified and examined. These forms were analyzed in relation to their meaning, grammatical function, phonetic representation, and interaction with diacritic marks. The resulting inventory distinguishes between lexical and non-lexical units. Of the total, 8,678 forms were classified as lexical, while 2,870 forms were classified as non-lexical.

This distinction is important because not every written monosyllabic form should automatically be considered a word. Some forms have clear lexical or grammatical functions, while others exist primarily because the writing system permits particular combinations of letters and diacritics. Although such non-lexical forms may not appear as ordinary dictionary words, they remain relevant to the description of the language because they help identify the range of possible sound and writing combinations.

Thus, the purpose of the present study is not simply to count monosyllables. Rather, it is to understand the different linguistic roles performed by monosyllabic forms and to document their relationship with the wider Lepcha linguistic system.

Objectives of the Study

The study has three principal objectives.

First, it aims to identify and classify the monosyllabic units found in Lepcha according to their linguistic function and semantic status. This involves determining whether individual forms function as independent words, morphological components, grammatical markers, or other meaningful units.

Second, the study examines the relationship between monosyllabic forms and diacritic marks. Diacritics may affect the written and phonetic representation of a form, and certain combinations may produce pronounceable forms without producing standard lexical or grammatical words. Understanding these combinations is therefore important for describing the relationship between orthography and linguistic structure.

Third, the study seeks to establish a comprehensive inventory that distinguishes meaningful lexical units from non-lexical forms. Such an inventory can provide a useful foundation for future studies of Lepcha pronunciation, morphology, spelling, dictionary-making, teaching materials, digital language resources, and language preservation.

Background and Conceptual Framework

The study of monosyllabic forms involves an examination of the smallest units through which meaning and sound are organized in a language. A central principle of the present research is

that a written form cannot be understood solely on the basis of its appearance. Its linguistic status depends on how it functions within the language.

For this reason, the study recognizes two major categories: lexical units and non-lexical units.

Lexical units are forms that possess identifiable meaning or linguistic function. They may occur independently as words or may operate as components of larger words. Some can function as prefixes or suffixes, while others can serve marking or sound-related functions. Their importance therefore extends beyond the traditional idea of vocabulary.

Non-lexical units are forms that do not possess an established independent meaning or function as ordinary words. Many of these forms result from combinations of letters and diacritics that represent possible sounds. Although they may not occur as conventional lexical items, they remain relevant because they demonstrate what is possible within the writing and sound system.

The conceptual framework consequently treats language as an interaction between meaning, grammar, sound, morphology, and writing. A form that appears unusual from a lexical perspective may nevertheless have value for phonetic or orthographic description. Similarly, a monosyllabic form that appears small in structure may have an important morphological role when it is attached to another form.

Methodology

Data Collection

The study began by constructing a comprehensive inventory of monosyllabic units in the Lepcha script. The inventory contains 11,548 forms. Each form was examined to determine its linguistic status and its relationship to the sound and writing system.

The analysis considered whether a form functioned as an established word, a meaningful morphological element, a grammatical or marking unit, or primarily as a representation of a possible sound. Particular attention was given to the effects of diacritic marks because these can produce different written and phonetic forms.

The methodology did not restrict the inventory to established dictionary words. Forms were retained when they represented pronounceable or structurally possible combinations within the system. This approach allowed the study to document both actual lexical resources and broader orthographic or phonetic possibilities.

Classification Criteria

The forms were classified into two broad groups.

Lexical units include forms that carry meaning, function as words, or perform a clear grammatical, morphological, or marking role.

Non-lexical units include forms that do not have an established independent meaning but represent possible sound combinations or arise from particular combinations of letters and diacritics.

The principal criterion was therefore linguistic function rather than simple written possibility. A form could be written in the Lepcha script without necessarily qualifying as a lexical word. The distinction between the two categories makes it possible to describe the language more precisely.

Analytical Procedure

After compiling the inventory, each form was examined according to its lexical, morphological, grammatical, and phonetic characteristics. The analysis asked whether the form could function independently as a word, whether it could operate as a morpheme such as a prefix or suffix, or whether it mainly represented a possible sound.

Forms that could be pronounced but did not satisfy the conditions for lexical or grammatical status were retained as non-lexical units. This procedure resulted in 8,678 lexical forms and 2,870 non-lexical forms.

Results and Analysis

Distribution of Monosyllabic Units

The inventory contains 11,548 monosyllabic forms. Of these, 8,678, approximately 75%, are lexical units, while 2,870, approximately 25%, are non-lexical units.

Category	Proportion	Number of Forms	Approximate
Lexical units		8,678	75%
Non-lexical units		2,870	25%
Total		11,548	100%

The distribution demonstrates that the majority of the identified monosyllabic forms have a recognized linguistic meaning or function. At the same time, the substantial number of non-lexical forms shows that the sound and writing systems cannot be adequately described through lexical vocabulary alone.

Lexical Monosyllabic Units

The 8,678 lexical forms represent the meaningful component of the inventory. Some function as independent words and therefore constitute part of the basic vocabulary of the language. Others occur as parts of larger words and contribute to word formation.

Monosyllabic units can therefore operate at different levels. A form may have independent lexical status in one context but participate in a larger morphological construction in another. Prefixes and suffixes are particularly significant because they demonstrate how small linguistic elements can modify or contribute to the meaning and structure of larger forms.

Some lexical units also have marking or sound-related functions. This demonstrates that lexical status is not limited to forms that operate as ordinary nouns, verbs, adjectives, or other

conventional word classes. The inventory reflects a broader understanding of linguistic function.

The findings consequently show that monosyllabicity does not determine linguistic importance. A single-syllable form can be a complete word, a grammatical component, a morphological building block, or a marking element.

Non-Lexical Monosyllabic Units

The 2,870 non-lexical forms do not function as established words, but their inclusion is important for understanding the structure of Lepcha. These forms arise largely through combinations of letters and diacritics that create possible sound forms.

If such forms were excluded, the resulting inventory would provide only a partial account of the language's sound and orthographic possibilities. Non-lexical forms therefore serve as evidence of the difference between what can be represented by the writing system and what is actually established as a lexical item.

Their documentation may also be useful for future research. Changes in pronunciation, the introduction of borrowed vocabulary, or the development of new linguistic forms may make previously unused combinations relevant. Maintaining these forms in the inventory therefore provides a broader record of the language's structural possibilities.

Phonetic Function of Non-Lexical Forms

One of the important contributions of non-lexical forms is their value for phonetic documentation. A vocabulary list containing only meaningful words cannot fully represent all possible sound combinations within a language.

The non-lexical inventory helps identify forms that are pronounceable or structurally possible even though they do not correspond to established words. These forms can therefore provide information about the boundaries of the sound system and the relationship between sound and written representation.

This is particularly important in language documentation. A comprehensive linguistic description should not necessarily exclude unusual or unused forms simply because they lack independent meaning. Instead, they can be recorded separately and identified according to their status. Such documentation provides researchers with a more complete picture of the language.

Diacritic Marks and Grammatical Constraints

Diacritic marks constitute one of the central issues identified by the study. The presence of a diacritic does not automatically guarantee that the resulting form is a legitimate word or grammatical construction. Some combinations may be writable and pronounceable but may not satisfy the grammatical or lexical conventions of the language.

This observation creates an important distinction between orthographic possibility and linguistic acceptability. The writing system may allow a combination to be represented, while the grammar or vocabulary of the language may not recognize that combination as a standard word.

For this reason, some forms are classified as non-lexical. They remain part of the inventory because they demonstrate what can be produced through the interaction of letters and diacritics, even though they do not function as established words.

The study therefore suggests that the relationship between diacritics and grammar deserves further investigation. A more detailed future study could examine the grammatical restrictions governing these combinations and determine how orthographic rules correspond to actual linguistic usage.

Morphological Significance

Monosyllabic units are also significant from a morphological perspective. They may function as basic lexical items or become components of larger words through processes of word formation.

Prefixes and suffixes demonstrate this particularly clearly. A monosyllabic unit can attach to another form and influence its meaning, grammatical function, or written structure. Consequently, the study of monosyllables contributes to an understanding of how larger linguistic forms are constructed.

The present study does not attempt to provide a detailed analysis of every prefix and suffix. Its primary purpose is to identify and classify monosyllabic forms. Nevertheless, the inventory provides a useful starting point for future morphological research. Further studies can examine individual affixes, their distribution, their meanings, and their relationship with particular categories of words.

Lexicographic and Documentation Significance

The inventory has important implications for lexicography and language documentation. The lexical component can contribute to the development and refinement of Lepcha dictionaries by identifying meaningful monosyllabic forms and distinguishing them from structural possibilities.

At the same time, non-lexical forms can help explain combinations that may otherwise appear confusing to learners or researchers. Keeping lexical and non-lexical forms separately documented can make language databases more precise and useful.

Such classification can also support digital language resources. A language database that distinguishes established vocabulary from possible phonetic or orthographic forms can provide a clearer foundation for future computational and educational applications.

The study therefore has value beyond theoretical linguistics. It contributes to the practical documentation of Lepcha and can support future dictionary development, language teaching, spelling studies, and preservation initiatives.

Discussion

The findings demonstrate that Lepcha monosyllabic forms perform a wider range of functions than a conventional word list might suggest. The identification of more than eleven thousand

forms indicates the importance of monosyllabic structures within the broader organization of the language.

The distinction between lexical and non-lexical forms is particularly significant. The 8,678 lexical units demonstrate the extent to which monosyllables contribute directly to meaning, vocabulary, morphology, and grammatical functions. The 2,870 non-lexical units, meanwhile, demonstrate that the structure of the language cannot be understood solely through established vocabulary.

The inclusion of non-lexical forms also highlights the importance of documenting the boundaries of a writing system. There is a difference between what can be written, what can be pronounced, and what functions as an accepted word. The inventory captures these distinctions and therefore provides a more comprehensive description.

Another important issue is the relationship between diacritics and grammatical constraints. Some combinations produced through diacritic usage may be structurally possible but not lexically or grammatically established. This suggests an area for future research into Lepcha orthographic conventions and grammatical rules.

The morphological significance of monosyllabic units is equally important. Their occurrence as prefixes, suffixes, and other components shows that monosyllabic structure is connected directly to word formation. Understanding these units can therefore help explain how larger words are constructed and how meanings are modified.

Overall, the inventory-based approach adopted in the study provides a broader picture of Lepcha than a conventional dictionary alone could provide. A dictionary necessarily emphasizes established lexical items, whereas the present inventory also records structural and phonetic possibilities.

Implications for Lepcha Linguistics and Language Preservation

The findings have several implications for Lepcha linguistic research and preservation.

First, the inventory provides a systematic reference for scholars and future generations interested in the structure of the language. By recording both lexical and non-lexical forms, the study preserves information that might otherwise remain undocumented.

Second, the distinction between lexical and non-lexical forms can contribute to improved dictionaries and language databases. Clearly identifying which forms have established meaning and which represent structural possibilities can make future resources more accurate and useful.

Third, the inventory may contribute to language teaching. Learners can benefit from a clearer distinction between established words and forms that are merely possible within the writing system. Such classification can also help explain pronunciation and spelling patterns.

Fourth, the documentation of non-lexical forms contributes to long-term preservation. Language systems change over time, and forms that are not currently used as words may become relevant in future linguistic developments. Recording them creates a more complete historical and structural record.

Finally, the study highlights the importance of continued linguistic research. Further work can investigate individual monosyllabic forms in greater detail, examine prefix and suffix systems, analyze the grammatical restrictions associated with diacritics, and develop more comprehensive lexical and digital resources.

Conclusion

This study provides a comprehensive inventory and classification of Lepcha monosyllabic forms. The identification of 11,548 forms, divided into 8,678 lexical units and 2,870 non-lexical units, demonstrates the complexity and importance of monosyllabic structure in Lepcha.

The study shows that monosyllabic units cannot be treated simply as isolated words. They may function independently as lexical items, operate as prefixes or suffixes, contribute to word formation, perform marking functions, or represent possible sounds within the writing system. The distinction between lexical and non-lexical units therefore provides an important analytical framework for understanding Lepcha.

The role of diacritics is another major finding. Not every combination of letters and diacritics produces a standard lexical or grammatical form. Some combinations are possible within the writing system but remain non-lexical. Recording these forms nevertheless contributes to a more complete description of Lepcha phonetic and orthographic possibilities.

The inventory also has practical importance. It can serve as a foundation for dictionary development, language teaching, grammatical research, orthographic analysis, digital language resources, and preservation efforts. By documenting both meaningful forms and structural possibilities, the study creates a broader record of the language.

The present work should therefore be regarded as a foundation for further research rather than a final description of all aspects of Lepcha monosyllabic structure. Future studies can investigate the grammatical behavior of prefixes and suffixes in greater detail, examine the rules governing diacritic combinations, and explore the relationship between written forms, pronunciation, and actual usage.

References

- Diringer, David. Introduction to the Lepcha Script and Graded Lessons to Read the Language. India Office Library, n.d.
- Lepcha, Ajuk. International Lepcha Language and Literary Seminar. The Indigenous Lepcha Tribal Association, Mani Printing Press, 2009.
- Lepcha, Gideon. Learning Lepcha in the 21st Century: An Introduction to the Wangboo Framework, the Modern Way to Learn the Ancient Language. 2022.
- Lepcha, Ugen. Lepcha Verse: Monosyllabic Musings. Amit Offset Press, 2024.
- Mainwaring, G. B. A Grammar of the Rong (Lepcha) Language. Mani Printing Press, 1876.
- Shipmoo, Ugen. A Lepcha to English Homonym. Sikkim Lepcha Literary Organization, n.d.

- Sitling, G. T. Lepcha Primer 1. Mayel Lyang Lepcha Development Board, Mani Printing Press, 2015.
- Tamsang, K. P. The Lepcha-English Encyclopedic Dictionary. The Lepcha Association, Mani Printing Press, 1982, 2002.
- Tamsang, Lyangsong. International Lepcha Language and Literary Seminar. The Indigenous Lepcha Tribal Association, Mani Printing Press, 2009.
- Tamsang, Lyangsong. Raong Ring Hlaap. Mani Printing Press, 2012.
- Tamsang, Norbu Tshering. New Lepcha-English-Nepali Dictionary. Sikkim Lepcha Literary Organization, 2005.
- Thomas, P. C. Lepcha Antonyms. 2018.