

Abstract 24

A Corpus-Based Morphological Analysis of Sinhala Verbs

W. S. N. Dilshani¹, Prof. Gihan Dias²

¹Department of Computer Science and Engineering, University of Moratuwa
nimashadilshani@cse.mrt.ac.lk

²Department of Computer Science and Engineering, University of Moratuwa
gihan@uom.lk

Verbs are essential components of a meaningful sentence and are important in understanding the sentence structure. This paper presents a morphological analysis of Sinhala verbs by combining traditional Sinhala grammar with an analysis of current usage based on a corpus of official documents. Sinhala verbs may be classified into a number of groups based on their morphology. However, there is currently no well-defined methodology to classify a particular verb. It is hypothesized that verb morphology patterns may be identified by analysis of a Sinhala text corpus. On the basis of that hypothesis, this research proposes a classification for Sinhala verbs based on their morphology which allows the morphological analysis of verbs in Sinhala text, and also the derivation of morphological rules for each class of verbs. This classification and rules are derived from an analysis of the corpus of official documents. Additionally, the rules were tested by applying them to another part of the corpus. This also allows the identification of irregular verbs, which do not fall into standard classes. With the analysis, it was identified that the usage of tenses in contemporary official documents is more complex than those given in grammar texts and different combinations of Sinhala grammatical forms are used to denote the time periods among standard tenses. Moreover, other writing forms of Sinhala were identified and it is shown that the existing classification of verbs in traditional grammar is insufficient to handle modern usage of the language.

Key words: Inflection, Morphology, Sinhala, Traditional Grammar, Verbs