Abstract

Resource development mainly focuses on well-described languages with a large amount of speakers. However, smaller languages may also profit from language resources which can then be used in applications such as electronic dictionaries or computer-assisted language learning materials. The development of resources for such languages may face various challenges. Often, not enough data is available for a successful statistical approach and the methods developed for other languages may not be suitable for this specific language. This report describes the morphological analysis on the Bishnupriya Manipuri language using finite state transducers. As of now, there is no computational work available for this language. Since finite state morphology is one of the successful approach applied in a wide variety of languages over the years, we adapted this approach to analyze the morphology of the Bishnupriya Manipuri language. For this purpose, a corpus also has been developed in the Bishnupriya Manipuri Language. The text used in the corpus is collected from web blogs, web sites, Bishnupriya Manipuri version of the Wikipedia and some typed text. We have used the xfst tool on this corpus for generating the finite states.

Keywords: *the Bishnupriya Manipuri language, morphological analysis, corpus, finite state transducer*