DEVELOPING AN ELECTRONIC MOBILE DICTIONARY FOR TAMIL: A CORPUS - BASED STUDY

Dr. Thennarasu, S, Department of Linguistics, Central University of Kerala
Mr. Winston Cruz, Chief Resource Person in NTM, Central Institute of Indian Languages
Dr. R. Prabagaran, Lecturer in Tamil, BBMP Composite Junior College for Girls

Abstract

This study aims to discuss the issues in developing an electronic mobile dictionary based on Tamil Corpus of Singapore. This is a unique dictionary that we present here for Tamil with three layers. The mobile based dictionary that we develop would be very much useful for the learners, teachers and researchers of Tamil. This paper also discusses some of the issues that rose during the compilation of the frequency based dictionary.

Keywords: Mobile Dictionary, Corpus-Based, Tamil, Language Learning and Teaching

1. Introduction

The aim of this paper is to develop 50k words mobile dictionary for Singapore Tamils. Singapore is one of the countries in Southeast Asia which has Tamil, English, Malay and Chinese as official languages. English is the most common language in Singapore so all people speak, read and write English but Tamils in Singapore are losing their knowledge of reading and writing and there is a need to improve their knowledge of Tamil. So the present mobile dictionary that we develop for Tamil language will certainly be a support and help for the learners of Tamil in Singapore in particular and all over the world in general. Sedelow, Sally Yeates (1985) in his articles on Computational Lexicography quotes Alan Walker Read notes that "the compilation of a dictionary is lexicography; lexicology is a branch of linguistics in which, with the utmost scientific rigor, the theories that lexicographers make use of in the solution of their problems are developed" (Read, 1983).

Lexicons are now being given considerable attention both by pure and applied linguists. As the technology has advanced, there is assumption about the value of automated dictionaries, which would enable students to access them from terminals not only for traditional help such as spelling and definitions, but also for enlightened word games intended to enhance "sensitivity to alternative meanings of words" (Miller 1979). Lexical databases provide more flexible and rapid access to much larger bodies of words, computational linguists are giving more attention to this aspect of their discipline.

According to Juan-Pedro and Rica-Peromingo (2012) Lexicography deals with the meaning and use of words. In previous decades, lexicographers have investigated the meanings of words and synonyms, but current lexicographic research has been extended using corpus-based techniques to study the way that words are used and, in particular, how lexical associations are used. Lexicography is, therefore directly connected to phraseology because both disciplines study sets of fixed expressions (idioms, phrasal verbs etc.) and other types of multi-word lexical units.