False Friends in Tamil - Malayalam Languages

S. Thennarasu

Introduction

Traditionally, words have not been regarded by language teachers in classrooms. However, in the last decades of the 20th century, linguists and researchers started to show a growing interest in the role of lexical acquisition in language teaching which has lasted up to the present (Nation, 1990; Lewis, 1993; Singleton, 1999; Schmitt, 2000; Bogaards and Laufer, 2004; M Luisa Roca-Varela 2015). In fact, vocabulary is now considered as a central issue for language learners to be effective and accurate in the communication process. The present study acknowledges the prominent role of word knowledge while analyzing the inter-linguistic phenomenon of false friends. False friends are lexical items in different languages that resemble each other in form but have different meanings (Chalker and Weiner, 1996; Colman, 2009).

Despite the fact that false friends have a long tradition in language research, these words are a current issue for those learning and working with languages since they can be found in many contexts of our real life but also in linguistics.

Tamil-Malayalam

1) *mati* means 1.estimate 2.value(sth.); respect, 3.knowledge/moon in Tamil but whereas *mati* in Malayalam means 'enough'.

2) *camcāram* in Tamil means 'wife' but whereas *samcāram* in Malayalam means 'speech' If a Malayali says that "(uṇṭē) samcāramsariyilla" which means 'Your speech is not good/true'. Whereas in Tamil "samcāramsariyilla" which means '(your) wife is not good'.

Singapur-Tamilnadu dialects

3) *cāram* means 'veesti' in Singapore dialect but whereas *cāram* in Tamilnadu dialect means 'a kind of stare case used for construction purpose' (Renugadevi , 2011).

Tamil-Kannada

4) *maadu* means 'do' in Kannada whereas *maadu* in Tamil means 'cow'; My experience with the word "*maadu*", when I was working with LDC-IL, CIIL, Mysore, asked my colleague (who was a programmer) to work with me for developing a tool then he said "niimaadunaamaadilla" which mean 'You do and I am not doing' but whereas the literal meaning of that in Tamil means "You are cow and I am not".

Tamil-Telugu

5) *nalla* in Tamil means good but whereas *nalla* in Telugu means 'black'. If we say "maṇappeṇnallāirukkiṟāḷ," bride is looking good. In Telugu, it would be "pellikūthruṇalla-gaunthi" which means 'bride is black'.Therefore, if we look into the senses of these words in pair of languages, we observe that the previous examples are just a symbolic representation of the thousands of situations in which false friends may occur and produce misunderstandings among speakers with different language backgrounds. Such language problems should be detected and avoided. As a contribution, the present paper seeks to identify the linguistic problems having to do with only Tamil-Malayalam false friends.

TEACHING OF AFFIXES ON VOCABULARY LEARNING OF TAMIL

S. Thennarasu

Abstract

The present study is deal with most frequently occurring affixes and their importance on vocabulary learning and teaching of Tamil students. Vocabulary is a fundamental basis of every sentence in every language. Students' vocabulary knowledge can be advanced through a variety of techniques. One of the most popular strategies in vocabulary learning supported by many researchers is using affix knowledge (Fotokian, M. & Rahmatipasand Z. 2015) which can assist learners in learning unfamiliar words. Therefore, this study investigates different affixes for vocabulary learning of Tamil based on a 20 million words of Tamil corpus.

(**Keywords:** affixes, learning, teaching, corpus, vocabulary)

1.0 INTRODUCTION

The trends of linguistic theory play a role in the development of language pedagogy. The earlier dominance of Chomskyan School of Linguistics partly accounts for the general neglect of vocabulary acquisition in favour of syntactic development. However, the fact that foreign language learners generally see vocabulary learning as their first priority and report that they encounter considerable difficulty in vocabulary learning is extensively recognized by language teachers and repeatedly pointed

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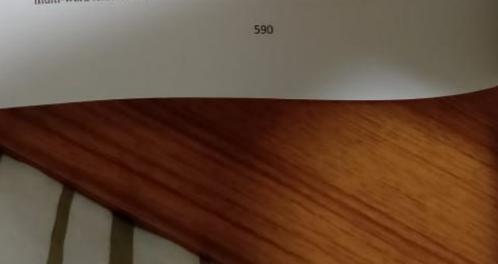
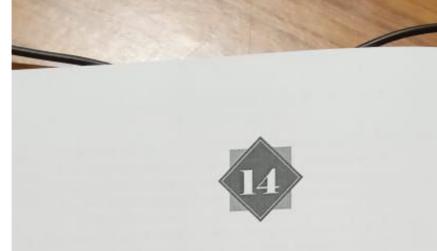*Keywords: Mobile Dictionary, Corpus-Based, Tamil, Language Learning and Teaching*

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.

<div align="center">**14**</div>

Collocation in CIIL Corpus: Implications for Language Teaching and NLP

<div align="right">*Thennarasu, S.*</div>

1. Introduction

Collocations are a feature of natural languages that are not well addressed by current Tamil language teaching and current models used for NLP. Language is full of word combinations that occur more frequently than expected (Joachim Wagner, 2008). The scholar Graeme Kennedy has stated that Palmer (1933) is one of the most influential English language teaching specialists of the 20th century, who adopted the term *collocation* for recurring groups of words. He defined a collocation as "a succession of two or more words that must be learned as an integral whole and not pieced together from its component parts" (e.g. Tamil. *talaimai aluvalakam* 'chief office', *mālai vēḷai* 'evening time' etc.). Palmer went so far as to suggest that even a "selection of common collocations... exceeds by far the popular estimate of the number of single words contained in an everyday vocabulary." The possibility that there are many more collocations to learn than there are words in a language perhaps helps explain why learning a language usually takes so long in comparison with other complex learning tasks.

Palmer's (1933) pioneering work on collocations in English language teaching was paralleled in different branches of the language sciences. Among a number of scholars who took account of the phenomenon of collocation, Firth (1957) emphasized the importance of both linguistic collocation and situational context for the description of languages in his maxim, "You shall know a word by the company it keeps".

Corpus-based evidence was used by Sinclair (1991) to support what he called the *idiom principle* in language learning and use (characterized by the use of routinized combinations of words in speech and writing), and to highlight the neglect of collocations in the theory and practice of English language teaching.