DEVELOPING ELECTRONIC DICTIONARY FOR ANTONYMS AND SYNONYMS IN TAMIL

Dr.p.vijaya, CAS in Linguistics, Annamalai University. Chidambaram

Introduction

This paper is mainly concerned about developing electronic dictionary for Antonyms and Synonyms in Tamil. This e-dictionary makes learning simple and it covers the lexical items used in Tamil. This widespread introduction of new lexical items would be helpful to the students and research scholars who work in the Tamil Language and Linguistics.

Lexicography means the collection of lexical items and descriptions of the way they are used. A dictionary is a simple tool that helps us to pronounce in the language. It gives meanings to words which cannot be understood or words that have more than one meaning. Besides meaning, it also provides information on the reader syllables, intonation and pronunciation of words. Dictionaries are considered practical sources of information on which learners and teachers can depend. Dictionaries could take different formats like traditional print dictionaries of varies types, handheld electronic dictionaries, concordance, indexes, terminologies, online dictionaries and CD ROM dictionaries depending on the purpose of their required usage.

Handheld electronic dictionaries, also known as "pocket electronic dictionaries" or PEDs, resemble miniature clamshell laptop computers, complete with full keyboards and LCD screens. Because they are intended to be fully portable, the dictionaries are battery-powered and made with durable casing material. Dictionaries have become very important tools for learning of a language especially the new forms of electronic dictionaries. Several technological developments have led to the invention of electronic dictionaries in the operational processes of teaching and learning process. Language can be studied from the point of view of language structure and language use. Study of language structure is called structural or formal linguistic study. Study of language use is called as functional linguistic study.

Languages are described qualitatively in terms of grammatical units like nouns, verbs, noun phrases, verb phrases, subject, object, agent, goal, etc. to explain the structure of the language. Most of the grammar units take a sentence as the minimum unit for the description of the structure. The structure has been studied from different viewpoints and many linguistic theories have emerged to account for the syntactic pattern of the sentences. Nowadays, more and more students use electronic dictionaries instead of ordinary dictionaries. The discussion about using electronic dictionaries has been popular in recent years. The most important advantage of an electronic dictionary is that it's very convenient to use. Whenever you meet new words or expressions, you can know the meaning quickly.

Electronic Dictionary

An electronic dictionary is a dictionary whose data exists in digital form and can be accessed through a number of different media. An electronic dictionary can be used to refer to any reference material stored in electronic form that gives information about the spelling, meaning, or use of words. It is the retrieval system, rather than the information content, which

makes electronic dictionary use such a revolutionary experience compared to the consultation of a hard-copy dictionary.

Online Dictionaries and Electronic Dictionaries

There are different kinds of websites available for Tamil Dictionaries which can be utilized for learning Tamil language. For example, there are several types of online dictionary including:

- www.sol.com.sg/classroom/dictionaries/html is a website for accessing the Tamil lexicon.
- website for an online web based English Tamil dictionary is www.geocities.com/Athens/Acropolis/8780/. The University of Chicago has developed a website to access the Cre-A: Dictionary of Contemporary Tamil (on line version) - www.lib.Uchicago.edu/LibInfo/Subjects/SouthAsia/.
- Another website for Tamil- English, English-Tamil and Tamil-Tamil Dictionary www.murasu.com/akaram will be useful for learning Tamil language through online.
- They typically offer monolingual and bilingual dictionaries, one or more thesauruses, and technical or specialized dictionaries. Examples include The FreeDictionary.com and Dictionary.com.
- 'Premium' dictionaries available on subscription, such as the *Oxford English Dictionary.* Dictionaries from a single publisher, free to the user and supported by advertising. Examples include *Collins Online Dictionary*, Duden Online, Larousse bilingual dictionaries.

Some online dictionaries are regularly updated, keeping abreast of language change. Many have additional content, such as blogs and features on new words.

An electronic dictionary is a dictionary whose data exists in digital form and can be accessed through a number of different media. Electronic dictionaries can be found in several forms, including:

- dedicated handheld devices
- apps on smart phones and tablet computers or computer software
- a function built into an E-reader
- CD-ROMs and DVD-ROMs, typically packaged with a printed dictionary, to be installed on the user's own computer as free or paid-for online products

Computational Lexicography

It covers computational methods tools designed to assist the various lexicographical tasks, that includes preparing lexicographical evidence from many sources, recording the database form of the relevant linguistic inform data editing of lexicographic entities, and dissemination of lexicographical products.

The computational lexicography involves the following processes:

- using computers to assist ID, capture, encoding, dissemination of lexicographic information.
- Implementing/ leveraging/ integrating lexicographic resources in computational tasks.

- supplying/ comparing/ evaluating MRDs (Machine Readable Dictionaries).
- tools for all of above.

Main sources of information for Computational Linguistics work.
MRD (Machine Readable Dictionaries)

1. In the past, the interaction between computational linguistics and lexicographers has been virtually limited to building sample tools for producing indexes and in concordance of large corpora.
2. lexicographers should use the techniques developed by computational linguistics for browsing machine readable dictionary in order to reuse their data.
3. future electronic publishing benefits end user by improving access to the wealth of information currently available in dictionaries but not easily retrievable via alphabetical order alone.

The processing of text corpora
- ➢ one of the reasons is that computational linguistics, which covers both linguistics and computers, develops very fast so that what is held true today may be superseded tomorrow
- ➢ Another obvious reason is that computational linguistics is so vast a domain that a single individual cannot be expected to cover the whole of it
- ➢

What do we mean by lexicographical information?

A text could be defined as a set of one or more (written or spoken) words forming a consistent whole from a semantic, syntactic, and pragmatic point of view.

Lexicographer makes use of 'Primary' texts and of 'secondary' texts.

Texts---➔   Processing ---➔   Lexicographical information

By 'lexicographic information' we mean the kind of information one can expect to find in a dictionary.

Synonym

Synonyms are words that have the same or very similar meaning. The absolute identity or similarity in the meaning of several lexemes is called synonymy. The different lexemes that are related to one meaning are called synonyms. The meaning of these words need not be exactly the same. Most words can have one or more synonyms.

The synonyms can be subgrouped into personal nouns and abstract nouns. *al:va:r*, *irai, iraivan, puranta:r, mannavan, ve:ntan, ve:ntar* and *ve:ntu* are the synonyms of the personal noun, *aracan* 'king'. Similarly *Ilukkam, Ilukku, ili, e:tam, ta:, ti:tu, pali, ma:cu* and *vatu* are the synonyms of the abstract noun *kurram* 'offence' .

Antonym

Antonyms are words which have almost opposite meanings. Most words can have one or more antonyms. The opposite of meaning within two-member sets of lexemes is called antonym and the sets of two words which refer to the opposite extreme from identity of meaning, are called binary contrasts or antonyms. There are several binary contrasts in the Tamil Language. As in the case of synonyms, the antonyms can also be sub-grouped into personal nouns, abstract nouns, adjectives and verbs. *periyar/ciriyar, va:ymai/poymai, eliya/ariya* and *a:rum/a:ra:te:* are the binary contrasts of personal noun, abstract noun, adjective and verb respectively.

Aim and Objectives

The main aim of this research paper is to study the various synonyms and antonyms used in Tamil Language with their respective references and finally create an inventory of synonyms and antonyms found in Tamil Language.

- For the proposed work, about 500 antonyms and synonyms in Tamil have been collected from tirukkural books and school text books.
- To develop an e-dictionary for antonyms and synonyms in Tamil.
- To provide an asset to Tamil language and also its community like students, researchers, teachers, and so on.
- To make the students enrich their vocabulary power through a technology-incorporated dictionary i.e. a virtual software (programme) written in a programming language (i.e. JAVA).
- To investigate the design, construction and use of electronic dictionaries in natural language processing the notion of re-usability of lexical resources is taken as a focus, enabling considerations of existing lexical resources, including publishers (MRDs) and lexical database, as well as the constructions of new resources.

Scope of this study:

- The dictionary of Tamil is a completely new one of its kind for students of linguistics.
- This dictionary covers the terms used in General and Applied Linguistics.
- This widespread introduction of new terms would be the great help to the students and Research scholars of in Tamil/ Linguistics.

Methodology

The process of the research article has four different phases' , viz. data collection, the structure of a data entry, analysis and verification & evaluation.

a. Data collection

The lexical items related to the same meaning (synonyms) and opposite meaning (antonyms) have been collected along with the references of their respective and different sources like Tamil text books, Journals, English Glossaries, Language corpus and website.

b. The structure of a Data Entry

- The collected data have been digitalized and alphabetized in Tamil.

- Tamil entry: The "head word"
- In Tamil script

In Synonyms

murd; (1)	Nte;jd; (382>549>899)
	Nte;J (551)
	Gue;jhh; (780)
	,iw (547>563)
	,iwtd; (690)
	kd;dtd; (553)
	Nte;jh; (481)
	Ms;thh; (447)
nry;tk;	nry;tk; (125>247>755)
	Mf;fk; (112>562)
	cilik (89>558)
	cud; (1263)
	gw;W (606)
	NgW (61)
	nghUs; (1009)
	khL (168)
	jpU (179>215>482>616>408>374)

In Antonyms

,ir (238)	tir (238>239>230)
nry;yplk;	my;yplk; (301)
M$o;	Ngh$o; (371)

c. Analysis

The data collected have been linguistically analysed and categorically identified as personal nouns, abstract nouns, adjectives, verbs and so on. Finally, all the lexical items collected and analysed in different lists category-wise.

d. Verification

The verification phase is very important one as it enables the product of the research to attain perfection and accuracy without any error. So the whole developing e-dictionary for antonyms and synonyms work have been meticulously verified and neatly in relevant categories.

e. Evaluation

The final stage of the process of the research is a scrupulous evaluation of the work, which would be undertaken after the accomplishment of full-fledged work. The evaluation would

be in the point of view as to what extent the product of the project will cater to needs of the educated Tamil public in general, and the students and research scholars in particular.

f. Computation of data

Finally, all the information has been instructed to the computer by using the JAVA programming languages for compilation of the e-dictionary for Antonyms and Synonyms in Tamil.

The advantages of using electronic dictionaries

Let's have a look at its advantages.

Electronic Dictionaries are user-friendly, very fast and easier to carry.

- Students relate better to it and are more enthusiastic and sounds can be heard.
- They are better for class-based activities, while paper dictionaries would be better for homework based tasks
- What's more, electronic dictionaries are becoming more and more advanced: they can pronounce the words clearly, provide sample sentences to illustrate word usage and store difficult words for special memorization.
- You don't need to waste much time turning the page seeking for the new words.
- That's special function may help you remember some new words with high efficiency.
- Last but not least, the designing of electronic dictionaries are more and more portable, and they are easy for students to carry.
- When learners use an electronic dictionary their word search becomes faster.
- This allows learners to spend more time on reading comprehension or on searching word meanings.
- A comparison between electronic and paper dictionary revealed that a lot of differences exist between them in some features.
- The vocabulary items covered in the electronic dictionary exceeds the limited number in the paper one.
- Some features do not exist in the paper dictionary such as databank, reference book, voice recorder, calculator, MP3. In addition, it does not include speech features, updates and interactive learning functions like irregular verbs, idioms, dialogues, sentence structure, accent correction, and grammar explanations.
- Obviously, the electronic dictionary is faster, lighter and more mobile than the paper based dictionaries.

Developing Tools (Screen Shot):
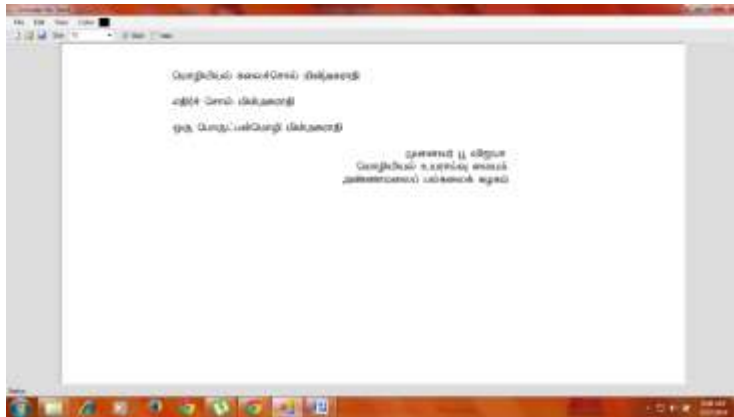
Conclusion:

Finally the research article has been brought out in the shape of e-dictionary for Antonyms and Synonyms in Tamil. The electronic dictionaries are one of the ways used in Tamil computing. This is because lexical items in Tamil and their various functions and roles of their thematic expressions are computed by using various programming languages. This can be, for example, of great benefit to learners of Tamil as a second language.

References:
1. Atkins, S. & Rundell, M. *The Oxford Guide to Practical Lexicography*, Oxford University Press 2008: 238-246 language forums
2. Chen, Yuzhen, 'Dictionary use and EFL learning: a contrastive study of pocket electronic dictionaries and paper dictionaries', in *International Journal of Lexicography* 23 (3), 2010:275-306 Franklin MWS-1840
3. De Schryver, Gilles-Maurice, 'Lexicographers' dreams in the electronic dictionary age', in *International Journal of Lexicography*, 16(2), 2003:143-199
4. Lew, Robert. 'Online Dictionaries of English' in Fuertes-Olivera, Pedro A. and Henning Bergenholtz (eds.), *E-Lexicography: The Internet, Digital Initiatives and Lexicography*. London/New York: Continuum, 2011: 230-250.
5. Raja, S (2012) Role of electronic dictionaries in Tamil Language Teaching and Learning, International Confeence Infit, Research article, Annamalai University.
6. Tiberius, C. and Niestadt, J. 'The ANW: an online Dutch dictionary', in Dykstra, A. and Schoonheim, T. (eds), *Proceedings of the XIV Euralex Congress*, Leewarden, 2010: 747-753
7. Trap-Jensen, L., 'Access to Multiple Lexical Resources at a Stroke: Integrating Dictionary, Corpus and Wordnet Data', In Sylviane Granger, Magali Paquot (eds.), *eLexicography in the 21st Century: New Challenges, New Applications* Louvain-la-Neuve: Presses universitaires de Louvain, 2010:295-302.