8.15 The Structure of a Dialect Dictionary of Agricultural Vocabulary in Tamil

Abstract—In the present scenario of science and technology advancements several indigenous occupational vocabularies are being replaced gradually by English. This often leads to replacement of culturally related vocabulary like occupational items with textual hybridity. In this context, this paper mainly focuses on various issues related to textual hybridity in Tamil agricultural vocabulary. Apart from this, this paper shows the significance of dialect surveys and the compilation of a dialect dictionary. The significance of dialect surveys is to bring out the variant forms of different regions. Recent research in the field of metalexicography, based on the structure of the dictionaries identified by Hausmann and Wiegand (1989), and Weigand (1989), are made use of in this study.

I. INTRODUCTION

This paper shows the significance of dialect surveys and the compilation of a dialect dictionary. The significance of dialect surveys is to bring out the variant forms of different regions. Recent research in the field of metalexicography has focused on the structure of the dictionaries identified by Hausmann and Wiegand (1989) and Weigand (1989).

The proposed dictionary will examine the different structures of a dictionary the macrostructure. Here the macrostructure is the selection of lemma sign in the spoken variety only. The macrostructure of the proposed dictionary has two important structures; the information structure is not a lexicographic definition but the lexical description given by the farmers in Tamil. Another important structure in the macrostructure is providing the variation found in the different dialect regions. This is the main aim of the proposed dictionary. For this, the microstructure may be different when compared with the microstructure of a general monolingual dictionary. Here all the variant forms are clearly marked by different labels. These variants are treated in the macrostructure as dialect variant form (DVF).

Lexicography is an applied field, the theoretical background for which is provided by linguistics. Almost all the branches of linguistics provide information for the making of a dictionary. Phonetics and phonology are useful for providing information about the pronunciation standardization of the script and alterations in the form of headwords. Grammar (morphology and syntax) is essential for providing grammatical information of different types, like parts of speech of words, gender distinction and causal relation of nouns, conjugation of verbs, etc. Dialectology, sociolinguistics, psycholinguistics, etc. are useful in giving usage and other types of labels to the headwords. But the branch of linguistics which is most important for the compilation of dictionaries is semantics, especially lexical semantics, as meaning is the central and vital component in the structure of a dictionary entry. This paper mainly focuses on the method of preparation of an article or dictionary entry for agricultural vocabulary in Tamil.

II. DICTIONARY DEFINITION

According to the Oxford English Dictionary (OED), the word 'dictionary' is derived from the Latin form 'dictionarius', used in 1225 by English poet and grammarian Joannes de Garlandia (John of Garland) as the 'title of a collection of Latin vocables' and the word 'dictionarium' is used about a century later by Peter Bechorius (or Becharius). The first recorded appearance of the word dictionary as such is dated to 1526 by the OED; later the word was used by Thomas Elyot in 1538. The French word dictionnaire seems to have been used for the first time by Robert Estienne in 1539 (Bejoint, 2004).

In Webster’s third new international dictionary (W3) the definition of dictionary is slightly more detailed, but very similar: 'A reference book containing words usually alphabetically arranged along with information about their forms, pronunciation, functions, etymologies, meanings, and syntactic and idiomatic use.'

Matore (1968) finds all definitions of dictionary in dictionaries less than satisfactory. The only one that satisfies him is the one that is proposed by the OED. It runs: 'A book dealing with the individual words of a language (or certain specified classes of them), so as to set forth their orthography, pronunciation, signification and use, their synonyms, derivation and history, or at least some of these facts: for convenience of reference, the words are arranged in some stated order, now, in most languages, alphabetical; and in larger dictionaries the information given is illustrated by questions from literature'.

Zgusta (1971) uses slightly more precise formulation. 'A dictionary is a systematically arranged list of socialized linguistic forms compiled from the speech-habits of a given speech-community and commented on by the author in such a way that a qualified reader understands the meaning of each separate form, and is informed of the relevant facts concerning the function of that form in its community.'
III. DICTIONARY TYPOLOGY

Zgusta (1971) opined that when a lexicographer sets out to compile a dictionary, he has to take two basic decisions: (1) what part of the total vocabulary of a language the proposed dictionary will cover, and (2) what type the proposed dictionary will belong. Both the aspects of typological classification of dictionaries are quite useful in understanding the classification and locating the proposed dictionary. Dictionaries can be classified into various types on the basis of different criteria. Zgusta classifies dictionaries into two major divisions namely, linguistic and non-linguistic dictionaries. The linguistic dictionaries are concerned with the words or lexical units of languages, called word books. The non-linguistic dictionaries are not concerned with the words, but with realia or denotata (thing) and they are called encyclopedias, or thing books.

Classification of linguistic dictionaries has been attempted by a number of scholars such as Chherba (1940), Sebok (1962), Maleki (1967), Comyn (1967), Gelb (1968), Zgusta (1971), Al Kasimi (1977), Singh (1982), and Svensen (1993). Contrasting features of dictionary typology are divided into two basic classes, viz., internal features and external features. The internal features are concerned with the aspects of the nature of language such as paradigmatic vs. syntagmatic contrast, aspects of form/meaning, time and area of vocabulary covered. External features are concerned with the target group or user of the dictionary for whom the dictionary is compiled i.e., size, purpose, arrangement of entries or words, number of languages, etc. Internal features are theoretical in nature, whereas the external features are applicational or concerned with the use or practical utility of the dictionary.

IV. METHODOLOGY

Agricultural vocabulary have been collected from fifty points comprising of the five major dialect regions of Tamil Nadu, namely, (1) Eastern dialect, (2) Western dialect, (3) Northern dialect, (4) Southern dialect, and (5) Central dialect. The distance between each point was twenty-five to thirty kilometers distance. Informants were selected from densely populated areas. The informants are of different age groups and caste groups, both male and female. Old persons in the age group of fifty to seventy who have no educational background were also selected. Lexical items relating to various agricultural areas like soil, irrigation, water lifting instruments, weather and season, ploughing, leveling the field, seeds, crops, harvesting, manure, cattle, etc. were included in the detailed questionnaire which was prepared for this purpose. The questionnaire contained 527 direct questions. In addition to these, a number of additional questions were also asked to collect sufficient data. The data were recorded in a tape-recorder. Data were analyzed and presented in the form of a dictionary format.

V. DESCRIPTION OF LEXICAL MEANING

Zgusta (1971) says that there are four basic instruments which are used for the description of lexical meanings. They are: (1) The lexicographic definition, (2) synonyms, (3) exemplification, and (4) glosses and labels. A number of other theoreticians include synonyms as a kind of definition (Svensen, 1993; Kipfer, 1984) Landau, 1989; Hartman and James, 1998). What Zgusta calls exemplification, others such as Landau (1989) refer to as illustrative quotations, Ison (1986) as illustrative phrase, Svensen (1993) as examples of usage and Kipfer (1984) calls illustration or exemplification used as the organic part of the definition within parenthesis.

Label is another instrument used to describe the lexical meaning, especially in the compilation of a dialect dictionary (proposed dictionary) and in connotation. Zgusta regards labels as a species of glosses. But others such as Svensen (1993) discuss it separately, while Landau (1989) and Kipfer (1984) describe it as a part of usage.

VI. LEXICOGRAPHIC DEFINITION

There are different types of lexicographic definitions (1) Paraphrase, (2) Referential/analytical/intentional definition, (3) extensional definition, (4) formulative definition, and (5) definition by synonyms. The second type, what Zgusta (1971) calls lexicographic definition, Ison (1986) calls as referential or analytical definition, while Svensen (1993) calls intentional definition, which are modified versions of lexical definitions ‘genus proximum’ and ‘differentia specifica’. The concept represented by the headword is called the definiendum and the definition (verbal description) is called the definiens. Intention denotes the content of the concept which is defined as the combination of distinctive features or what Zgusta calls criterial features which the concept comprises. It expresses a generic conceptual relationship whereby concepts are arranged in classes according to similarities and differences noted between them. This has resulted in a hierarchical system with super-ordinate (hyperonym), sub-ordinate (hyponym), and co-ordinate (co-hyponym) concepts. The process of definition involves stating the superordinate concept next to definiendum, i.e., (genus proximum) together with at least one distinctive feature typical of the definiendum (differentia specificum) (Svensen, 1993). This type of
 VII. COMPONENT PARTS AND STRUCTURE OF A DICTIONARY

The present article focuses on the dictionary type in general and very particular to the structure of an entry. The word 'dictionary' has two meanings in a textural manner, i.e., (1) the whole book and (2) the word list, which constitutes the main part of the book. Both of them have two structures, the textural book structure and the textural word list structure, respectively. The word has several important units and structures. The basic unit of the dictionary is the treatment unit. The treatment units have a form and information relating to that form are brought together. The relation of form and information is that of topic and comment (Hausmann and Wiegand, 1989).

A form and information relating to a form are brought together under the addressing procedure. Each information item is addressed to a form called address. In any dictionary the most important item is the definition, but there may be hundreds of other information types, i.e., items. The most important address is lemma (headword or entry word), because the lemma belongs to the alphabetical access structure of the dictionary. Normally all the ordered set of lemmata of the dictionary forms the macrostructure. The lemma and the whole set of information items, which are addressed to the lemma, form the dictionary article. The macrostructure is the relative arrangement of the stock of lemmata in the word list (cf. Svensen 1993, Bengenholtz 1995). The macrostructure is the overall list structure which allows the compiler and the user to locate information in a reference work. “It is the macrostructure that determines under which lemma the lexicographical item is to be found” (Hausmann & Wiegand 1986). The macrostructure may have a single central word list or additional word lists. The arrangement of words in the macrostructure can be fully alphabetic or systematic (conceptual grouping) or a combination of both. The macrostructure of the proposed dictionary has a central word list. Normally speaking, the structure of information within the article is called the microstructure.

The proposed dictionary entry is given in the following format. The headword or article is given in Tamil script with bold letters, followed by the phonemic script (transliteration). The grammatical indication is also given and in addition to this, the English equivalent is also provided. After that, the lexicographic description is given in Tamil, which is added for better understanding of the target language group (Tamil language group). Many types of dictionaries will have the same structure but the proposed dialect has a different structure. That is, in addition to the lexicographic description, the variations found in the different regions are also presented in the same entry. Those variations are labeled. Such type of arrangement is called niching. It means a strict-alphabetical clustering of lemmata or articles that may or may not be semantically related. All the variant forms also find a place in the dictionary entry. There the description or definitions are cross-referenced. If there is no variation in any district to the headword, then it is believed that the speakers of particular district use the same form (i.e., the central region form). In such cases, no form is given in that entry. Since the central region (Thanjavur) has the standard spoken variety, the headword is given in that variety only. All the regions are labeled according to the nesting method. The sample entries are also given.
VIII. GUIDE TO THE INNER STRUCTURE OF THE PROPOSED DICTIONARY

1. Head word/entry word (in BOLD script)
2. Phonemic Transcription
3. Grammatical Item
4. English Equivalent
5. Definition
6. Semantic gloss
7. Lexical variation
8. Phonemic transcription of variant form
9. ~ cross-reference
10. Label Item
11. Sub-label Item

IX. CONCLUSION

Due to the influence of science and technology, many new technical terms in English come into existence in almost all fields of Tamil. Because of this reason, much native vocabulary of indigenous occupations are slowly getting replaced by English language. This may lead to permanent loss of the rich occupational vocabulary in Tamil. This study will be quite helpful in preserving the occupational vocabulary in Tamil and recording it in a dictionary format for future purpose. This study may provide a model for the future survey and to record the various occupational vocabularies in Tamil.
APPENDIX

SAMPLE ENTRIES OF THE TAMIL AGRICULTURAL DICTIONARY

அஹ்நையை: இந்தச் சொல்லுடன் அல்லாது இழைப்பியல் நீட்சி பெறுவது ஏனைய சொல்லுடன் இணையை செய்தது. (ூ)

காண்கியை: சிட்டுக்குரு என (யார் கூறும்-வாசிக்கிழை). (ூ)

காண்கியை: சண்ட்டுக்குரு என (சோதனை செய்யும்). (ூ)

காண்கியை: சண்செய்துரு என (இந்தச் சொல் குறிப்பிட்டு நீட்சி பெறுவது ஏனைய சொல்லுடன் இணைக்கும்). (ூ)

காண்கியை: சண்செய்துரு என (சயனை-செய்யும்). (ூ)

காண்கியை: சண்செய்துரு என (சோதனை செய்யும்). (ூ)

சான்று சண்ட்டுரு என 1) தாலையில் கமிட்டுவைத்தது போன்ற இழைப்பியல் நெட்சி பெறுவது ஏனைய சொல்லுடன் இணைக்கும். (சயனை-செய்யும்). (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) சேருது சண்ட்டுரு என (சயனை-செய்யும்). (ூ)

செய்துரு சண்ட்டுரு என 2) சோதனை செய்யும் என இழைப்பியல் நெட்சி பெறுவது ஏனைய சொல்லுடன் இணைக்கும். (சயனை-செய்யும்). (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) சேருது சண்ட்டுரு என (சயனை-செய்யும்). (ூ)

செய்துரு சண்ட்டுரு என 3) சோதனை செய்யும் என இழைப்பியல் நெட்சி பெறுவது ஏனைய சொல்லுடன் இணைக்கும். (சயனை-செய்யும்). (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) சேருது சண்ட்டுரு என (சயனை-செய்யும்). (ூ)

செய்துரு சண்ட்டுரு என 4) சோதனை செய்யும் என இழைப்பியல் நெட்சி பெறுவது ஏனைய சொல்லுடன் இணைக்கும். (சயனை-செய்யும்). (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ)

செய்துரு சண்ட்டுரு என 5) சோதனை செய்யும் என இழைப்பியல் நெட்சி பெறுவது ஏனைய சொல்லுடன் இணைக்கும். (சயனை-செய்யும்). (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) சேருது சண்ட்டுரு என (சயனை-செய்யும்). (ூ)

செய்துரு சண்ட்டுரு என 6) சோதனை செய்யும் என இழைப்பியல் நெட்சி பெறுவது ஏனைய சொல்லுடன் இணைக்கும். (சயனை-செய்யும்). (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) சேருது சண்ட்டுரு என (சயனை-செய்யும்). (ூ)

செய்துரு சண்ட்டுரு என 7) சோதனை செய்யும் என இழைப்பியல் நெட்சி பெறுவது ஏனைய சொல்லுடன் இணைக்கும். (சயனை-செய்யும்). (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) சேருது சண்ட்டுரு என (சயனை-செய்யும்). (ூ)

செய்துரு சண்ட்டுரு என 8) சோதனை செய்யும் என இழைப்பியல் நெட்சி பெறுவது ஏனைய சொல்லுடன் இணைக்கும். (சயனை-செய்யும்). (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) சேருது சண்ட்டுரு என (சயனை-செய்யும்). (ூ)

செய்துரு சண்ட்டுரு என 9) சோதனை செய்யும் என இழைப்பியல் நெட்சி பெறுவது ஏனைய சொல்லுடன் இணைக்கும். (சயனை-செய்யும்). (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) சேருது சண்ட்டுரு என (சயனை-செய்யும்). (ூ)

செய்துரு சண்ட்டுரு என 10) சோதனை செய்யும் என இழைப்பியல் நெட்சி பெறுவது ஏனைய சொல்லுடன் இணைக்கும். (சயனை-செய்யும்). (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) (ூ) சேருது சண்ட்டுரு என (சயனை-செய்யும). (ூ)
 TNM உலர் உலர் அல்லது உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உலர் உul 1}
REFERENCES