Chapter 1: Introduction

Language is the most effective medium for communication as it can be effortlessly employed to understand and express a wide range of thoughts and feelings. It is used to disseminate information to each other, to describe what we see around us, to reflect on our thoughts about ourselves, about each other and to share them with others. Using language for communication involves complex and intricate mental processes and an immense amount of knowledge about the meanings of words that carry the core information to be passed on. Each language user has a personal vocabulary store or the *mental lexicon*, from which they select the words for use and to which they refer the words they encounter in the utterances of others. The term mental lexicon refers to a language user’s mental or cognitive representation of words that allows inferring the referents of a word, the semantic categories to which a word belongs and/or the similarities in word meaning. This information stacked in the mental lexicon is an integral component of knowledge about the world present in the brain alternately referred to as *concepts*. Language is also instrumental in developing these concepts as it is relied upon to obtain insights about the world around us.

The mental lexicon, which is a part of conceptual knowledge, is assumed to consist of a large set of lexical entries for each word. Lexical entries refer to the information stored about a word that is essential to recognize, understand and differentiate that word from similar words. This information about the meaning of the words can be described in terms of semantic features. Semantic features are individual components of meaning which, when added together gives the complete meaning of the word. For example the word apple can be described using semantic features such as <fruit>, <red>, <juicy>, <sweet>, <grows on trees> etc. These features provide insight into the representation of the respective word, the concept corresponding to the word and categorization of concepts. Hence a variety of theories and models proposed to understand the semantic representation\(^1\) and semantic organization\(^2\) of language in the brain (e.g. Shallice, 1993; Jackendoff, 1990; Smith &

---

\(^1\) Semantic representation in the present study is viewed as studying the mapping of words in the mental lexicon to their respective concepts

\(^2\) Semantic organization in the present study is viewed as studying how words in the mental lexicon are grouped together into respective categories (e.g. category of animals)
Medin, 1981; Rosch & Mervis, 1975; Norman & Rumelhart, 1975; Minsky, 1975; Smith, Shoben, & Rips, 1974; Collins & Quillian, 1969) have considered the semantic representation in terms of semantic features. The semantic features are typically generated for a given set of concepts by asking the participants to list semantic features that they think are salient in describing respective concepts (E.g.: <animal>, <has four legs>, <barks> etc., for the target word ‘dog’). Even though the feature generation task uses words as stimuli, the semantic features nonetheless provide a window into conceptual representation as the word and conceptual knowledge are closely related. The stimuli that are used for generation of semantic feature can be nouns representing concrete concepts, abstract concepts to some extent, verbs representing actions and also adjectives. Factors such as familiarity and imageability of concepts and also frequency of occurrences of these concepts in a language play crucial role in generation of semantic feature.

Acknowledging the relevance of semantic features for understanding semantic representation in the mental lexicon and in formulating theories and models, it is true that collecting semantic feature production norms can provide a strong basis for research in Linguistics and Language sciences. Considering the importance of semantic features, extensive normative databases have been collected for English language (Garrard, Lambon Ralph, Hodges, & Patterson (2001) for 64 nouns; Randall, Moss, Rodd, Greer & Tyler (2004) for 193 nouns; McRae, Cree, Seidenberg & McNorgans (2005) for 725 nouns; Vinson & Vigliocco (2008) for 240 nouns and 216 verbs; Buchanan, Holmes, Teasley & Hutchison (2013) for 1,808 nouns, verbs, adjectives and other parts of speech; Devereux, Tyler, Geertzen & Randall (2013) for 866 nouns). Semantic features have also been collected in Dutch (Ruts, De Deyne, Ameel, Vanpaemel, Verbeemen & Storms (2004) for 338 nouns; De Deyne, Verheyen, Ameel, Vanpaemel, Dry, Voorspoels & Storms (2008) for 425 nouns) and in Italian languages (Kremer & Baroni (2011) for 50 nouns; Montefinese, Ambrosini, Fairfield & Mammarella (2012) for 120 nouns) and also from congenitally blind Italian participants (Lenci, Baroni, Cazzolli & Marotta (2013) for 50 nouns and 20 verbs).

In Indian languages, norms have been established for limited aspects of semantic components, restricted to the purpose of the particular study under consideration. Such norms have been established in Kannada (Karanth, 1984), Hindi
(Sharma, 1995), Malayalam (Asha, 1997), Telugu (Suhasini, 1997) for Linguistic Profile Test developed to assess language comprehension and expression. Ranganatha (1982) has established norms for relative frequency of phonemes and morphemes in Kannada. However, lexical semantic representation in adult speakers of Kannada, with particular reference to the semantic features has not been studied till date.

The collected semantic features are studied for regularities and patterns in the distribution of features as they contribute to a great extent to the understanding of semantic representation in the mental lexicon. The number of features generated by the participants for each target word is the most basic distributional analysis that is carried out. Featural weight is another variable that is found to be very useful which signifies the importance of each feature for a word based on participant’s discretion. The features generated are also classified into various types based on the information they carry such as visual, tactile or functional property of the word and are analyzed. The semantic features are then analyzed for distribution of each of these feature types as it is highly informative in explaining neural representation of concepts. The semantic features are also studied for featural correlation, which estimates the occurrence of one feature with respect to others. The distinctive features which help to distinguish between similar words and shared features that are relevant for many words are also studied. These featural properties may vary depending on the categories of target words, concreteness or abstractness of the words and frequency of occurrence of the word in a language.

The featural properties shed light on important aspects of nature of semantic representation in the mental lexicon as these statistical regularities form the organizational principles of various proposed semantic theories and models of meaning representation. The semantic features are also used to carry out accurate and quantitative testing of the claims about the structure of mental lexicon as proposed by these theories and models. Hence many theories of semantic representation such as prototype theory (Rosch & Mervis, 1975) and exemplar theories (Smith & Medin, 1981) are based on semantic features. Semantic features also form the basic ingredients of different kinds of models namely hierarchical network model of semantic memory and language processing (Collins & Loftus, 1975), Semantic Feature Comparison model (Smith, Shoben & Rips, 1974), Featural and Unitary Semantic Space (FUSS) model (Vigliocco, Vinson, Lewis &Garrett, 2004), vector
models of memory (Hintzman, 1986; Murdock, 1982), models of semantic computation (McRae, de Sa, & Seidenberg, 1997; McRae, Cree, Westmacott, & de Sa, 1999), object recognition (Plaut, 2002), word recognition (Harm & Seidenberg, 2004), and semantic memory (Hinton & Shallice, 1991; Plaut & Shallice, 1993).

The semantic features are also part of several semantic models which aims to demonstrate how particular patterns of semantic deficits is seen as a consequence of loss of different features caused by brain damage (Farah & McClelland, 1991; McRae et al., 1997; Devlin, Gonnerman, Andersen, & Seidenberg, 1998). This involves training artificial neural networks with input data obtained from the distribution analysis of feature properties that are predicted to be crucial. For instance, a model for words representing living and nonliving concepts constructed by Farah and McClelland (1991) is based on the evidence from the semantic features that the visual-perceptual features are predominant for living things whereas the functional features for nonliving things. In order to demonstrate the behavioural trends seen in patients with semantic deficits, the model was selectively lesioned by impairing either visual-perceptual or functional features. Models have been proposed based on featural correlation and distinctive features of living things and nonliving things to elucidate the progression of semantic deficits caused by Alzheimer's dementia (Devlin, Gonnerman, Andersen & Seidenberg, 1998). Thus the major purpose of collecting semantic features is to construct empirically derived conceptual representations that can be used to test theories of semantic representation and computation (McRae, Cree, Seidenberg & McNorgan, 2005). Semantic features also form the basis of many treatment strategies designed to treat comprehension deficits and anomia in persons with aphasia such as Semantic Feature Analysis (SFA) (Boyle & Coelho, 1995). Semantic feature based therapy techniques are also been evidenced to be effective in Bilingual persons with aphasia (Kiran & Roberts (2010) for Spanish-English bilinguals and French-English bilinguals; Rangamani & Prema, (personal communication) for Kannada-English bilinguals).

Studies involving semantic concepts and representation have been extensively researched in languages such as English, Dutch and Italian to name a few but there is a sparsity of research in Indian languages in the areas comprising of meaning representation, organization of words, models of semantics particularly in Kannada.
which differs in linguistic properties compared to English. Hence the present study was conceptualized to understand the structure of mental lexicon in Kannada.

The present research was designed to study lexical semantic representation and organization for a set of nouns and verbs in Kannada, by collecting semantic features generated for nouns and verbs from adult native speakers of Kannada. The semantic features obtained were subjected to analysis in order to assess the nature of distribution of different semantic featural properties. Specifically the analysis focused on evaluating the differences and similarities if any, in the distribution of featural properties across the domains of nouns and verbs. Further, the responses were analyzed for distribution of featural properties across different semantic categories to which the target words may belong. The study also attempts to develop a framework for a model of mental lexicon in Kannada based on the semantic featural properties obtained from the present study. Analysis will also be conducted in order to investigate differences and similarities if any in the distribution of featural properties between Kannada and English languages. The next chapter in this thesis presents a detailed review of literature summarizing the past research relevant to the current topic of study. Chapter 3 describes the methodology used for semantic feature collection, tabulation and construction of computer database of the generated semantic features. Chapter 4 reports the results of the analysis of featural properties carried out and description of the model generated followed by discussion of the same. A brief summary and conclusions derived from the study is also presented.