Department of English Language and Linguistics
University of Sheffield

The Effects of Semantic Clustering in L2 Word Learning: Evidence from an Action Research Study

Volume 1

Evagelia Papathanasiou

A Thesis
Submitted for the degree of
Doctor of Philosophy
October 2006

Abstract

In recent years, contradictory advice to teachers has been emerging from studies into the use of semantic links or networks in classroom materials and activities for vocabulary learning in a L2. There is some experimental evidence which suggests that learning semantically related words (e.g., body parts) at the same time makes learning more difficult (Tinkham, 1993, 1997; Waring, 1997; Finkbeiner, Nicol, 2003). There is also a theoretical framework that strongly supports the idea that it is very useful to present words of related meaning together so that learners can see the distinctions between them and gain a complete coverage of the defined area of meaning (Channell, 1981, 1990; Neuner, 1992; Dunbar, 1992). The following paradox appears: while the experimental evidence suggests that semantically related vocabulary does not help vocabulary learning, the EFL coursebook-writers present vocabulary in semantic clusters. The experimental evidence mainly derives from research using artificial language and not a natural L2. The purpose of our research is to investigate which of the two contrasting views will prove to be a useful tool in L2 vocabulary learning. The present study was influenced by action research. It was conducted in EFL classrooms with natural learners in Greece. The subjects were 31 intermediate EFL children and 32 beginners EFL adults. Two different ways of organizing new vocabulary for presentation were employed: a) presenting semantically related words (topic-related vocabulary i.e. *mugging*, *terrorism*, *forgery*, synonyms, antonyms or homonyms) together at the same time, and b) presenting vocabulary in an unrelated fashion (i.e. *carpenter*, *tornado*, *sage*). Short and long-term tests were administered to the students. The presentation will focus on the main conclusion that semantically related vocabulary impedes L2 vocabulary learning. Adult beginners performed significantly better on the unrelated vocabulary test compared to their performance on the related vocabulary test. Word frequency (in language) when combined with unrelated presentation of new L2 vocabulary appears to make a difference in students' performance.

Acknowledgments

I wish to acknowledge the guidance and support I received from several people in order to complete this thesis. First and foremost, I would like to express my gratitude to my supervisor Dr Gibson Ferguson, who has influenced my thinking on L2 vocabulary research. He provided me with useful advice, supervision, insightful comments and constructive criticism with suggestions and clarifications which guided my thinking on the issues investigated in this thesis. His advice on crucial statistical issues was invaluable.

Many thanks also go to members of the department of English Language and Linguistics whose help was of great importance to me. I am also grateful to Anastasios Melas and Yiannis Bikos who provided me with official permissions to conduct research in real L2 classrooms. I would also like to thank Prof. Achilea Papoutsis for all the help he has given me especially at the final stages of the study. I am particularly indebted to my close friends Angie Barry, Anita Zgrablitz, Argyro Pantelidou, Viki Tolli, Kelly Pycroft, and Katerina Dintsi for their continual moral support and encouragement during the course of my study.

The completion of this thesis would not have been possible without the love and moral support of my family (Alexandros, Panagiota, Andreas) and Chris who has been tolerant of my silent (and sometimes 'noisy') preoccupation with the study and my long hours on home computer; his continuing inspiration has helped me to discover what is important both in life and in the completion of this thesis. In acknowledgement of their role throughout the course of my Ph.D. research, I would like to thank all these people from the depth of my heart; this thesis is dedicated to them.

Evagelia Papathanasiou
Sheffield, October 2006

Table of Contents

CHAPTER SIX: Data Analysis and Discussion

CHAPTER SEVEN: Summary and Conclusions

List of Tables

List of Figures

List of Appendices

Preliminaries

The present study is an attempt to improve results found from previous research regarding L2 vocabulary presentation. The effect of related and unrelated manner of presentation has been the subject for various researchers. So far the existing studies have been tightly controlled either in terms of the use of artificial language or overlooking teaching procedure. This has contributed little in the application of results in a natural L2 environment. The study presented below tries to combine the use of natural L2 in a real classroom environment using a natural teaching procedure.

In recent years, contradictory advice to teachers has been emerging from studies into the use of semantic links or networks in classroom materials and activities for vocabulary learning in a L2. Arguments for the presentation of related vocabulary in sets are mainly based on theory and not on experimental evidence. Numerous writers (see Channell, 1981, 1990; Neuner, 1992) suggest teaching related words in sets. While Nation (2000) finds it a good idea to teach related words in sets, he also refers to a growing body of research which shows that it takes more time to learn words that relate to each other in groups than it takes to learn words that are unrelated to each other. Tinkham (1993) found that presenting L2 students with their new vocabulary grouped together in lexical sets of syntactically and semantically similar new words might actually impede rather than facilitate the learning of the words.

The following paradox appears: while the experimental evidence suggests that semantically related vocabulary does not help vocabulary learning, the EFL coursebook-writers present vocabulary in semantic clusters (topic-related vocabulary). The experimental evidence mainly derives from research using artificial language and not a natural L2 (Tinkham, 1993, 1997, Waring, 1997). Although Schneider, Healy and Bourne (1998) used French as a L2, their experiment was not applied to natural foreign language students in a L2 classroom neglecting actual teaching procedure.

The work reported in the present thesis was motivated by an underlying desire to investigate which of the two contrasting views would prove to be a useful tool in L2 vocabulary learning. In other words, the purpose is to discover whether the findings of

previous, conclusive research actually apply to a natural teaching environment using a real L2. The best way to do that is to conduct our study in a natural EFL classroom. If we want to enrich our understanding of language learning and teaching, we need to spend time looking in classrooms with natural learners. For this reason, the present research follows a plan that is partly action-research oriented.

For the purpose of this study 63 subjects were used in total. The first sample consists of 31 Greek EFL young learners (20 girls and 11 boys), intermediate level, and aged from 11 to 16 years old. The subjects in the second sample are 32 Greek adult beginners (22 female and 10 male), aged from 30 to 50 years old, who attend adult-classes (seminars) on English language. A total of 120 words were taught and tested (60 in semantically related group and 60 in semantically unrelated group). The subjects were taught the words by providing the L1 translation equivalent. This was done since it is a very common technique of teaching new vocabulary. Exercises were used during the teaching procedure in order to enhance learning. When tested, subjects were asked to write the L1 translation of the L2 word provided. This whole process tested receptive knowledge as the subjects were required to recall the word form and meaning by providing the Greek equivalent. Error analysis is not within the scope of the present thesis. The study was structured and organized in order to combine theory with practice and furthermore benefit the learner and not the researcher.

The present thesis starts with the existing literature around L2 vocabulary acquisition. The main interest will be focused on L2 vocabulary acquisition in relation to the L2 mental lexicon. What follows is what I believe is interesting to discuss and examine since L2 vocabulary acquisition has an extensive literature. The importance of vocabulary in L2 learning and teaching as well as the structure and organisation of the mental lexicon as a storage system will be central issues in the present thesis. There will be an examination of stages of lexical development and representation in L2. The interaction between L1 and L2 mental lexicon and the importance of form and meaning in the L2 vocabulary learning will also be discussed. In addition, L2 vocabulary learning strategies and teaching methods that attract a great deal of attention from L2 researchers will be analyzed. The present thesis is structured as follows:

- Chapter One presents the way the mental lexicon is operating. It examines the structure of the mental lexicon as a storage system and investigates the basic issues concerning lexical representation and development in the L2 mental lexicon.
- In Chapter Two current issues involved in the L2 vocabulary learning and teaching will be discussed. There will be an analysis of Nation's (2001) and Schmitt's (1997) taxonomy of vocabulary learning strategies as well as a presentation of explicit vocabulary teaching methods with a special attention to Sökmen (1997).
- Chapter Three examines and explains sense relations as important aspects of L2 vocabulary learning that facilitate depth of word processing and knowledge.
- In Chapter Four the two contrasting views of L2 vocabulary presentation are discussed: the first is that learning related words together at the same time makes learning more difficult, and the second is that presenting words of related meaning together makes learners see the distinctions between them and gain valuable knowledge of the defined area of meaning. The study continues with the actual research study and the data obtained by it.
- In Chapter Five there is the analysis of methodology used to conduct the present study.
- Chapter Six consists of the data findings which are thoroughly discussed.
- Finally, Chapter Seven is an overall conclusive summary which brings together the various topics presented in preceding chapters.

CHAPTER ONE

Mental Lexicon and Vocabulary Development in L2

1.1 Introduction

Chapter One starts with a few words on the importance of vocabulary in L2 learning and teaching. It examines the structure of the mental lexicon as a storage system and investigates the basic issues concerning lexical representation and development in the L2 mental lexicon. We will also look at some representative bilingual models of lexical processing. The interaction between L1 and L2 mental lexicons will be considered. Chapter One will also address the question whether the L2 mental lexicon as a storage system is more form-based than the L1 mental lexicon. There is evidence which supports the notion that meaning rather than form poses greater challenge in lexical acquisition in both L1 and L2, and that lexical items are increasingly processed in meaning, rather than form, as their integration into the mental lexicon progresses. It will also be suggested that vocabulary acquisition involves the building of connections between new words and words that already exist in the mental lexicon. The notions of lexical knowledge and lexical competence will also be discussed.

1.2 The Importance of Vocabulary in L2 Learning and Teaching

The studies of vocabulary acquisition and related areas of lexical research in L2 acquisition have been relatively neglected. This is commented on within the fields of language teaching and applied linguistics. Zimmerman (1997), for example, argues that the teaching and learning[1] of vocabulary have been undervalued in the field of second language acquisition (SLA). Compared to work in grammar, for instance, much less is known about the nature of the L2 learners' lexicon. Richards (1976) and Levenston (1979) criticized the neglect of vocabulary in favour of grammar and syntax. This stands in parallel with the fact that lexical errors are the most common among L2 learners[2].

---

[1] 'Learning' and 'acquisition' will be used interchangeably in this thesis; both will refer to the process by which knowledge is internalised. It is assumed that since all learning is to some extent cognitively controlled, the distinction between conscious (learning) and subconscious process (acquisition), is not one of kind, but of degree (see Laufer, 1997).

[2] See Meara (1984).

Both learners and native speakers view lexical errors as the most serious and disruptive obstacles to communication (Gass and Selinker, 2001:372).

During the last few decades, however, the area of vocabulary studies and research has not been neglected (Meara, 1987). The 1980s and 1990s have experienced a growing interest in vocabulary learning and teaching. In particular, there have been empirically based studies on the nature of the bilingual lexicon, vocabulary acquisition, and teaching. In order to minimize L2 lexical errors, researchers highlight the importance of vocabulary in language learning. McCarthy (1990) argues that communication in a L2 cannot happen (in a meaningful way) without words. Vermeer (1992) also claims that learning a new language mainly involves learning new words and that "knowing words is the key to understanding and being understood" (Vermeer, 1992:147).

1.3 Definition of the 'Bilingual Mental Lexicon'

Before we examine the nature of the bilingual mental lexicon, it is necessary to define its two 'constituents', the terms *bilingual* and *mental lexicon*. The term bilingualism has received many definitions. Li (2000) provides a long list of definitions which have been used to describe bilingual speakers taking into account factors such as age, proficiency, language status, socio-cultural aspects, etc. Bloomfield (1933:56) describes bilingualism as the "native-like control of two languages". Macnamara (1967:59-60), on the other hand, in rather looser terms, interprets bilinguals as "persons who possess at least one of the language skills [speaking, writing, listening and reading] even to a minimal degree in their second language".

Similarly, the term *(mental) lexicon* denotes the 'mental dictionary' (Aitchison, 1994) or else the permanent, long-term storage of lexical knowledge in the mind. Singleton (1999:15) provides the following definition and refers to what this lexical knowledge consists of.

> [The lexicon] *"constitutes that component of language or knowledge of a language which has to do with what one might call 'local' phenomena – the meanings of particular elements of a given language, the phonological and orthographic forms of these elements, and the specific ways in which they collocate and colligate".*

This means that the lexicon does not only include 'building blocks' of words (*elements*) but also the interrelations between and among words (or the grouping of words) based on their semantics, form and syntactic function[3].

Considering the above, in the present context, the *bilingual mental lexicon* is the 'mental dictionary' of a person who possesses two languages (regardless of proficiency level). As a result, there are the L1 lexicon (one that corresponds to a bilingual's lexical knowledge in their first language) and the L2 lexicon (one that corresponds to a bilingual's lexical knowledge in their second language). This suggests two major questions concerning the nature of the two lexicons. In what ways is the L2 lexicon different from or similar to L1? Is the L2 lexicon separate from or integrated with the L1 lexicon? Both questions will be discussed in the following sections starting with a special reference to word forms and word meanings (form and meaning mapping).

1.4 Word forms and word meanings

The underlying question about the human mental lexicon is to understand how words[4] are organized, accessed and represented in the mind. The first step is to investigate the relationship between word forms and word meanings, because the way that lexical choice (word form) and meaning are linked is central in the discussion of the domain of the lexicon. In the present section, I will explore some of the different ways in which lexical meaning has been approached by linguists.

Possible organisation of the linguistic sign, following Saussure[5] (1916), is depicted as a combination of a *signifier* (unit of expression or form) and a *signified* (unit of content or meaning) in the mental lexicon. The relation between signifier and signified is an age-old philosophical debate which cannot be undertaken here. The majority of studies of the bilingual lexicon adopted this binary representational format of a word as the mapping of form onto meaning. For this reason, the organization of the bilingual

---

[3] For a comprehensive discussion of these issues see Levelt (1989), Aitchison (1994) and Singleton (1999).

[4] The terms *word, lexical item and lexical unit* ('the union of a lexical form and a single sense', Cruse, 1986:24) will be used interchangeably in this thesis. It is acknowledged that lexical items in the lexicon can be multiwords (e.g. idioms or phrasal verbs). The present inquiry, however, is restricted to the study of word-sized constituents. A discussion of why the notion of linguistic unit is the more appropriate concept in the context of L2 vocabulary acquisition can be found in Bogaards (2001).

[5] For a critical discussion on Saussure's views of *signifier* and *signified,* see Lyons, 1977.

lexicon is referred to as 'lexicosemantic' (De Groot, 1995) organization, which means the representation and interrelation of word meaning and word form.

Therefore, the starting point of lexical semantics is the mapping between form and meaning. The fundamental question is whether words have a determinate meaning per se, or are mediated or supported by concepts as base units of linguistic information. The traditional view states that each lexical item (form) is associated with a concept, which in turn represents a referent in the 'real world'. This mapping is seen as being communicated through the medium of concepts (see Figure 1.1 from Ogden and Richards (1936, cited in Singleton, 1999:30)).

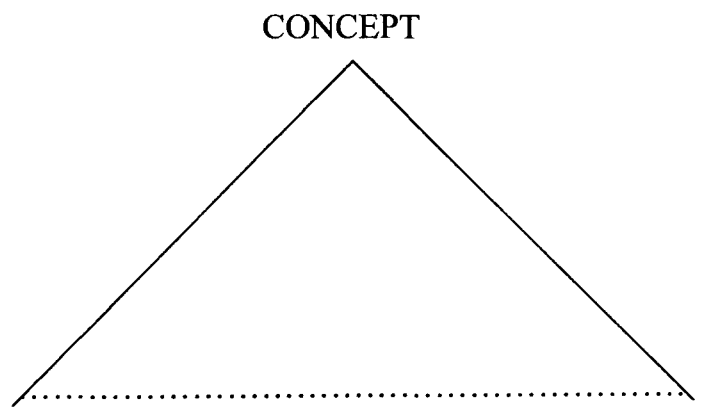


Figure 1.1 Ogden and Richard's (1936) basic triangle (simplified).

As Ijaz (1986) points out, words do not carry meanings by themselves, but only in relation to concepts. The traditional account of word meaning, based on Ogden and Richard's basic triangle, has been subject to various criticisms and has been rejected by various scholars for sometimes different reasons. For example, Lyons criticizes the ill-defined nature of 'concept': "As the term 'concept' is used by many writers, it is simply not clear what is meant by it[...]anyone who defines the meaning of a word to be the concept correlated with that word owes his readers some explanation of what kind of thing this concept might be" (Lyons, 1977:113).

Considering the above, the traditional view promotes the "atomistic view of language in which each word would be regarded as an isolated and self-contained unit" (Ullmann, 1962, cited in Singleton, 1999:30). This view was opposed by Structuralism generally taken to be rooted in Saussure's work, which emphasises the importance of system-

internal relationships, claiming that "linguistic units derive both their existence and their essence from their interrelations"[6] (Lyons, 1973, as cited in Singleton, 1999:30). In addition, the problems polysemy and synonymy[7] pose is another reason to reject the traditional view. Meanings of all words vary according to context, as a word is defined by its use in a wide range of contexts (Burgess and Lund, 1997). This means that words do not have simple meanings in terms of concepts, but rather that these 'meanings' are determined by their use. Lewis, in line with this view, rejects the temptation "to think of a word as having a fixed meaning, and to assume that words are in some sort of one-to-one isomorphic relationship with 'reality'" (Lewis, 1993:77).

1.5 Structure and organisation of the mental lexicon as a storage system

According to Schreuder and Weltens (1993), the central role of the mental lexicon for any model of language processing is that it functions as a bridge between the different constituents of a lexical item. This means that in the mental lexicon information from all different linguistic levels is combined. Phonology, orthography, syntax, argument structure, morphology and lexical semantics all appear in the entries of the mental lexicon (Schreuder and Weltens, 1993). The important thing is to understand how this information is represented and used.

This section considers the structure and organisation of the mental lexicon as a storage system and the ways in which that system is accessed under different conditions. In order to present what the lexicon as a storage system contains and how this system interacts with all aspects of language processing, I will start my discussion with Levelt's (1989) speech production model (Figure 1.2), which seeks to address all aspects of language processing and has inspired most L2 models on the organization of the mental lexicon. Singleton (1999) mentions that the primary perspective of the model is a productive one, although, receptive aspects of the processing are not entirely ignored.

---

[6] Relational Semantics is based on this approach, two of the most important strands being *semantic field theory* and *componential* or *feature analysis* (see Singleton, 1999).
[7] These lexical relations between words are discussed in detail in Chapter Three.

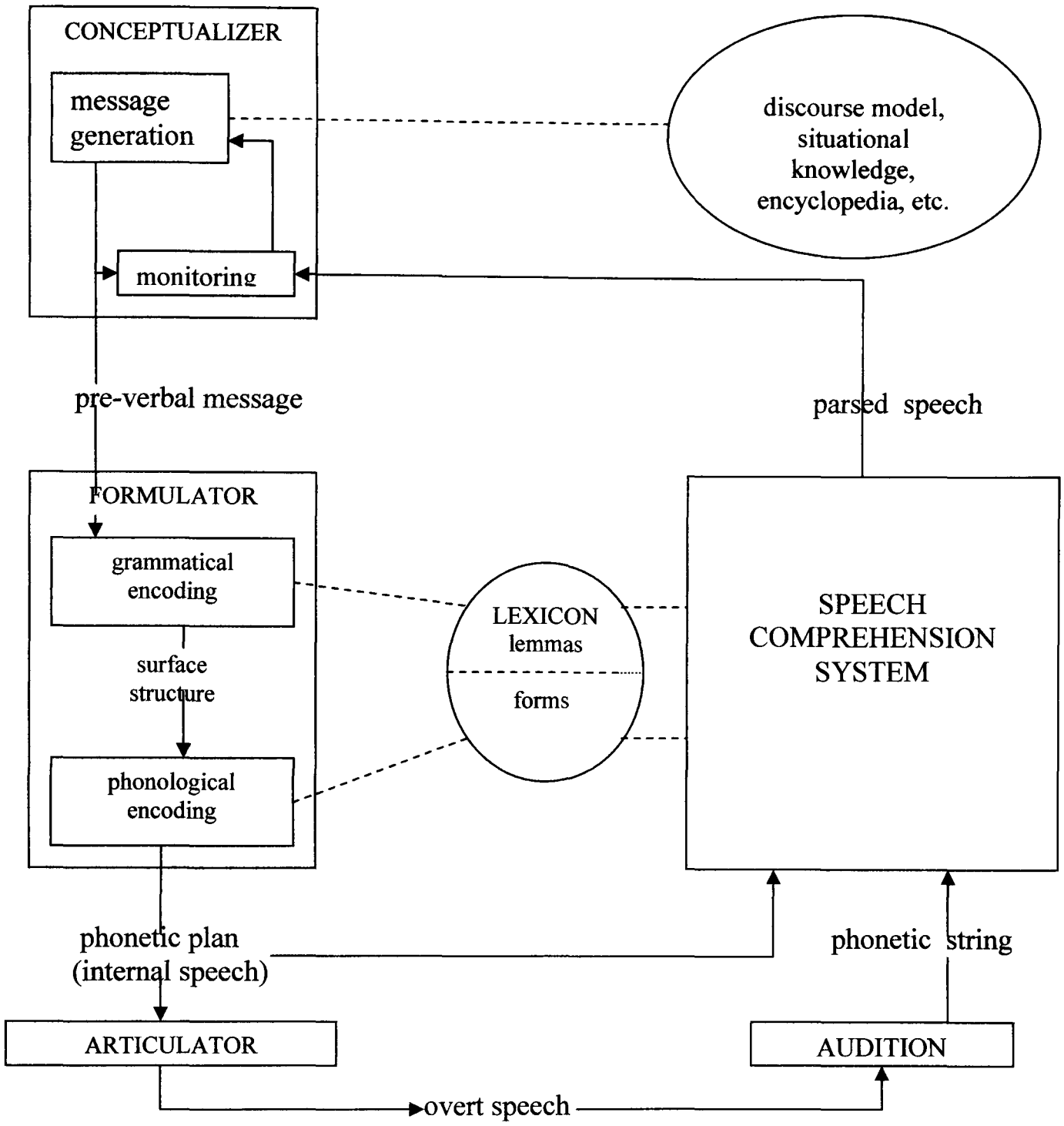


Figure 1.2 Levelt's blueprint for the speaker (based on Levelt, 1989:9).

According to Levelt (1989), there are two categories of component in this model, declarative and procedural. Declarative knowledge required for language processing includes general information about the world (encyclopedia), information about the specifics of particular situations (situational knowledge) and information about stylistic appropriacy relative to specific circumstances (discourse model). Declarative knowledge also includes lexical knowledge, both semantico-grammatical (lemmas) and morpho-phonological (forms[8]). The second component is called procedural knowledge and includes the Conceptualizer, responsible for message generation; the Formulator,

[8] Alternatively labelled *lexemes* in Levelt's terminology.

responsible for giving the pre-verbal message syntactic and phonological characteristics; the Articulator, responsible for executing overt speech emerging from the Formulator; the Audition component, responsible for analysing the speech signal into sound segments; and the Speech Comprehension System, responsible for making semantico-grammatical sense of phonetic strings received. The part played by the lexicon in speech production is seen by Levelt (1989:181) as absolutely central:

> *"This means that grammatical and phonological encodings are mediated by lexical entries. The preverbal message triggers lexical items into activity. The syntactic, morphological, and phonological properties of an activated lexical item trigger, in turn, the grammatical, morphological and phonological encoding procedures underlying the generation of an utterance."*

The lexical component is central to our interest here. L1 lexical entry in the mental lexicon is considered to contain semantic, syntactic, morphological and formal (phonological and orthographic) specifications about a lexical item (Jiang, 2000:48). More specifically, a lexical item consists of the lemma (semantic and syntactic information about a word, for example word meaning and part of speech) and the lexeme (morphological and formal information, for example, morphological variants of a word, spelling and pronunciation) (see Garrett, 1975; Levelt, 1989). Figure 1.3 provides a graphic description of a lexical entry (Jiang, 2000:48).

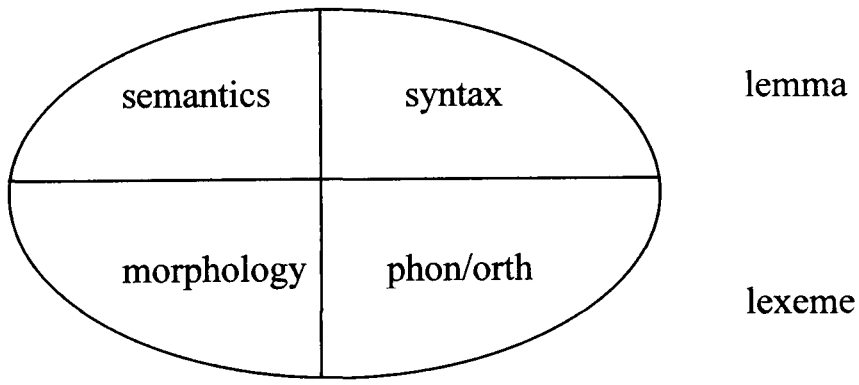


Figure 1.3 The internal structure of the lexical entry (adapted from Levelt, 1989).

One important aspect of the L1 lexical representation is that these different types of information are highly integrated within each entry and automatically become accessible (Jiang, 2000:49). This integration requires a high and extensive exposure to the language. L1 learners are able to extract semantic, syntactic and morphological information while becoming acquainted with the form of the word, because there is a highly contextualized input (Jiang, 2000).

1.5.1 Stages of lexical development and representation in L2

Questions related to the mapping process (mentioned earlier) have to do with the processes and mechanisms involved. According to Jiang (2002), what semantic information is used in the mapping process depends on one's theory of lexical and conceptual representations. If we consider meaning as an integral part of the information represented in lexical entries, as Levelt's (1989) model of lexical representation does, then one may ask what semantic information gets integrated in the L2 lexical entry. Let us examine this question by looking at the three stages of lexical development and representation in L2 provided by Jiang (2000).

According to Jiang (2000, based on Levelt, 1989), there are three stages of lexical development and representation in L2:

(a) At the initial stage, the use of a L2 word activates the links between L2 words and their L1 translations. In receptive use of the language, the recognition of a L2 word activates its L1 translation equivalent, whose semantic, syntactic, and morphological information then becomes available and assists comprehension. In productive L2 use, the pre-verbal message first activates the L1 lexical entry whose lemma matches the message fragments. The L1 word then activates the corresponding L2 words through the conscious recollection of L2-L1 connections established in learning the L2 word (see Figure 1.4).

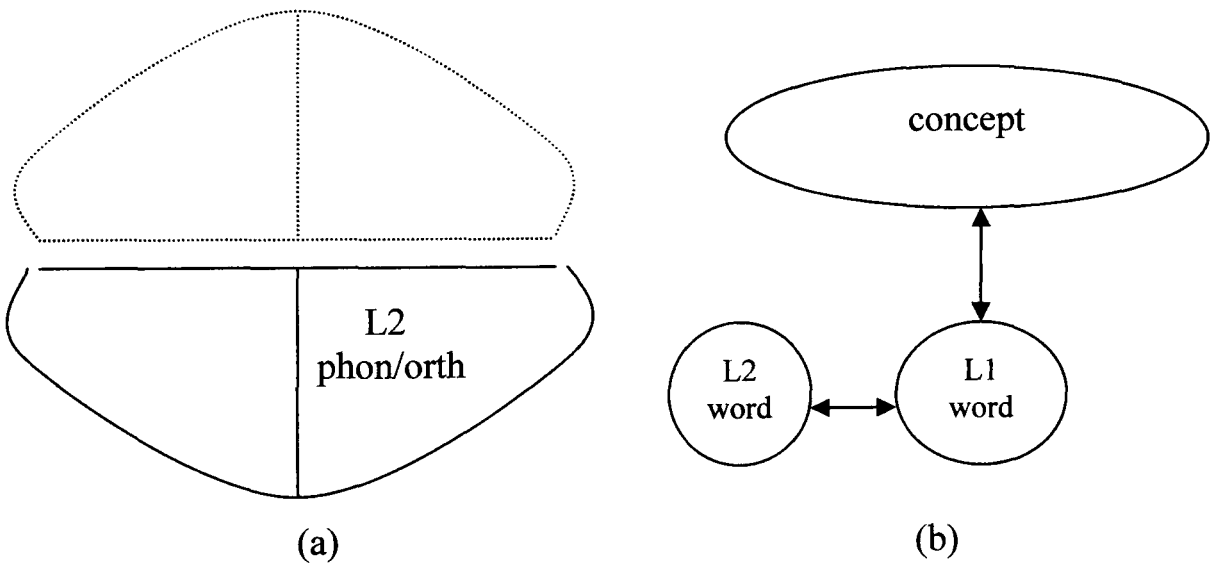


Figure 1.4 Lexical representation (a) and processing (b) at the initial stage of lexical development in L2.

(b) During the second stage, the L1 lemma mediation stage, as one's experience in L2 increases, the lemma space of a L2 word is occupied by the lemma information from its L1 translation and the L1 lemma information mediates L2 word processing. This means that information in L1 lemmas may be copied or attached to L2 lexical forms to form lexical entries that have L2 lexical forms but semantic and syntactic info of their L1 translation equivalents. In this way, stronger associations are developed between L2 words and their L1 translations (see Figure 1.5). But as Jiang points out, the representation of the information copied from the L1 lemma is weak, because part of the information is lost in transition.

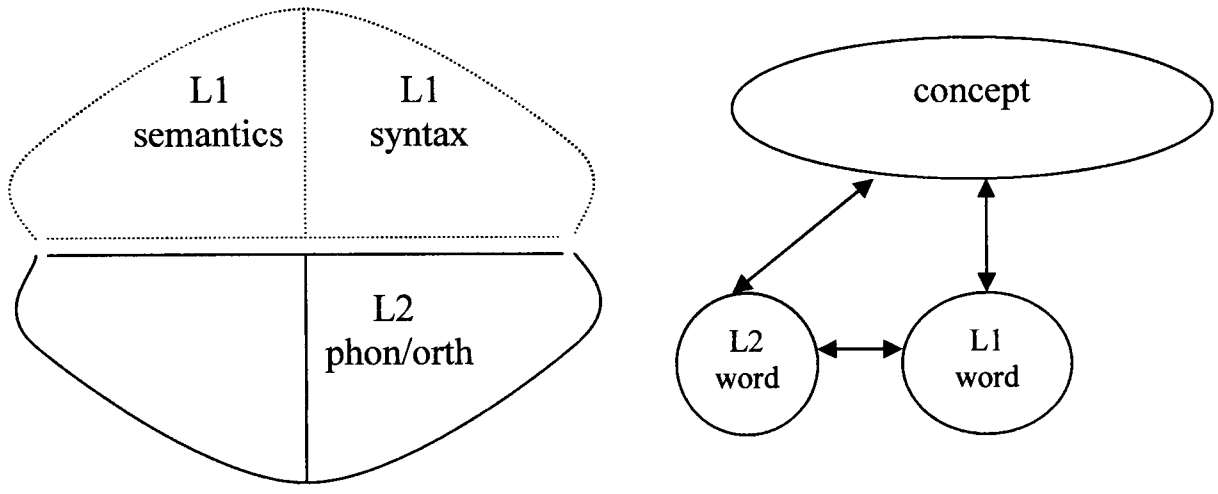


Figure 1.5 Lexical representation (a) and processing (b) in L2 at the second stage.

(c) At the third stage, the L2 integration stage, we have the full development of lexical competence, where a lexical entry in L2 will be very similar to a lexical entry in L1 in terms of both representation and processing (see Figure 1.6).

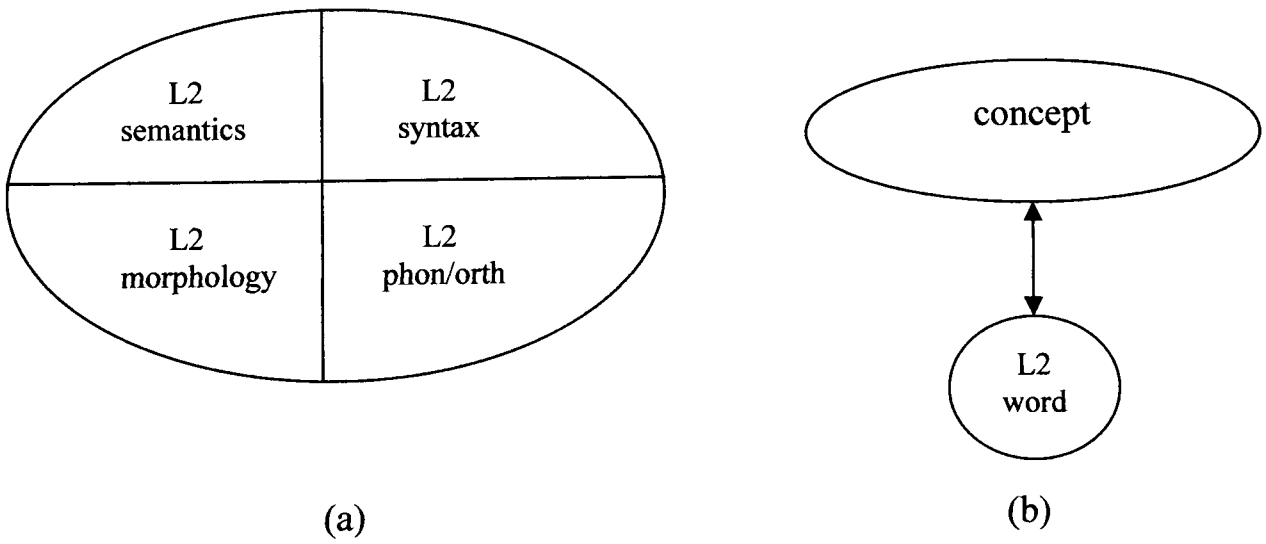


Figure 1.6 Lexical representation (a) and processing (b) in L2 at the third stage.

According to Jiang (2000), there are two factors which affect L2 lexical development process in the mental lexicon. The first one is that L2 learners often lack sufficient and highly contextualized input, and the second is that L2 learners may tend to rely on their already established L1 lexical and semantic system. In this sense, because the meanings of L2 words are understood through the L1 translations, the learner's language acquisition mechanism is less motivated to extract meaning from contextual cues (Jiang, 2000). The semantic information that is copied from the L1 translation stays in the L2 lexical entry and continues to mediate L2 word use even with continued exposure to the L2. As a result, even highly proficient L2 users will use L2 words on the basis of the semantic specifications of their L1 translations (Jiang, 2002). In other words, there is a continued L1 lemma mediation in L2 lexical performance.

A more simplified representation of lexical development in L2 is provided by the Revised Hierarchical Model (RHM), in which lexical and conceptual representations are treated as two separate levels, and no meaning is represented in lexical entries (e.g. Kroll & Stewart, 1994 and, Kroll & De Groot, 1997). In this case, the question becomes what concepts are mapped to L2 words. This hypothesis makes no claim regarding whether L2 words are linked to the existing concepts in the shared conceptual system or to new concepts in the system (Jiang, 2002).

According to Kroll & Stewart (1994), at an earlier stage of L2 language development there is lexical mediation whereby L1 translation equivalents are activated to facilitate access to concepts. Early in L2 learning, lexical form relations between L2 and L1 provide the basis of interlanguage connection. Over time, the patterns of conceptual activation will be correlated with L2 and the ability to conceptually mediate L2 directly will begin to develop. On this account, the conceptual information that is available initially for L2 will be the same as that available for L1. Only with increased opportunities to use the L2 word in different contexts will a distinct conceptual representation develop. Only with increased L2 proficiency are L2 learners able to access the meanings of L2 words directly. Words in each language (L1 and L2) are interconnected via lexical-level links and conceptual links (see Figure 1.7). The lexical-level links (connections) are stronger from L2 to L1 (solid line) than from L1 to L2

(dashed line) but the conceptual links (connections) are stronger for L1 (solid line) than for L2 (dashed line)[9].

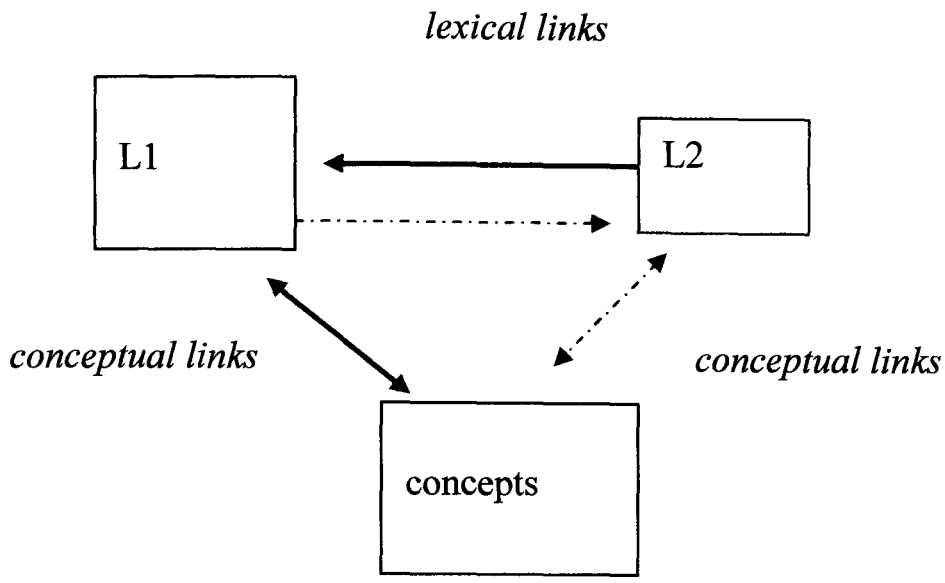


Figure 1.7 The Revised Hierarchical Model (adapted from Kroll & Stewart, 1994).

1.5.2 Interaction between L1 and L2 Mental Lexicon

In order to understand how vocabulary acquisition takes place, we have to understand how L1 and L2 words are stored in the mental lexicon. Hulstijn (1997:211) points out that there are four different hypotheses (for a review see De Bot, 1992):

1. The extended system hypothesis (L1 and L2 words are located in a single store)
2. The dual system hypothesis (L1 and L2 words are located in separate stores)
3. The tripartite hypothesis (similar words, such as cognates, are stored in a common store and language-specific words are stored in separate stores)
4. The subset hypothesis (although all words are located in a single store, the subset of L1 words are more strongly associated with each other than with those of the L2, as if L1 words and L2 words constituted two relatively separate 'families' within the entire 'community' of words)

The dichotomy presented by the first two patterns reflects the question of whether words from a bilingual's two languages are represented in a common, language-shared system or in two discrete, language-specific systems. There is a plethora of studies that

---

[9] The initial evidence for the revised hierarchical model came from translation experiments which showed that translation from L1 to L2 was slower, less accurate, and more likely to be influenced by semantic variables than translation from L2 to L1 (Kroll & Stewart, 1994).

set out to explicate the above question (for a review see De Groot, 1993, Kroll, 1993 and Kroll & De Groot, 1997)[10]. De Groot (1993), for example, does not refer to a single representational system for a given individual with more than one language at her or his disposal but rather to a mixed representational system. Most recent evidence though appears to favour the view that postulates that, depending on the level of representation[11], the two language systems can be both integrated and distinct. In particular, most research supports the view that the conceptual/lexical-semantic level is language-shared whereas the word-form level is language-specific. Gerard and Scarborough (1989) found evidence that while lexical retrieval is language-specific, suggesting a separate word-store for each language, lexical encoding and semantic/conceptual memory may be language-common. This leads us to the current question on the organisation of the bilingual lexicon which no longer concerns separate or common storage of L1 and L2 words, but at which levels and under what conditions L1 and L2 interact.

Before we present bilingual models of lexical processing, we have to look at the symbolic and connectionist paradigm in order to comprehend how the L2 mental lexicon is organized. According to Singleton (1999), the *Chomskyan/Fodorian* view of language is based on what is sometimes called the *Symbolic Paradigm*, the idea that cognition involves the manipulation of symbols. According to this view, the lexical development is viewed as an accumulation of entities. It envisions the lexicon as a dictionary, containing words subcategorized for a variety of syntactic, grammatical and semantic features. In addition to dictionary entries, the lexicon stores derivational, inflectional, and free grammatical morphemes.

The *Connectionist Paradigm*, on the other hand, does not interpret cognition as involving symbol manipulation but offers a different conception of the basic processing system in terms of connection strength rather than rules or patterns (Singleton, 1999). The term 'connectionism' relates to the fact that this approach derives from what is

---

[10] The study of the organisation of the bilingual lexicon has been tackled with a variety of paradigms, including word association, word recall, translation, picture naming in L1 and L2, etc. The existence of cross-linguistic effects in a bilingual task indicates a common, language-independent system in the bilingual memory whereas the absence of such effects suggests separate, language-specific systems.

[11] Bilingual models postulate either one level of linguistic representation or two levels (a lexical level, encoding word form, and a conceptual/semantic level, encoding word meaning) or three levels. For examples, see Kroll & de Groot, 1997.

known about neurophysiological activity in the brain. Aitchison (1992:31) points out that during any brain activity, numerous brain cells are active, sending out signals[12] to other neurons generating a 'network' of interconnected units. Connectionism assumes that the brain stores information in networks of nodes or relatively discrete knowledge structures. The concept of spreading activation allows us to conceive of the mental lexicon as a dynamic system. The process of creating form-meaning relationships occurs when neural networks are strengthened over time as the learner frequently encounters the item in the input. For example, two words may first be stored in an entirely unassociated fashion. As Hulstijn (1997) explains, later they may be linked via only one formal or semantic feature, and still later via more features. The strength of all these associations may differ, and the strength of each individual association may increase, and even decrease, over time (Hulstijn, 1997).

The concept of connectionism is clearly presented in Meara's (1997) model of L2 vocabulary acquisition, where an acquisition event (an unknown word in a text is somehow learned) consists of the building of a connection between a new word and a word that already exists in the learner's lexicon. Meara (1997) suggests that this connection might be a link between the new L2 word and its L1 translation equivalent, or it might be a link between the L2 word and an already known L2 word. In this sense, unknown words are words that have no connection of any kind to the lexicon, while known words are words that are connected to the lexicon. According to Meara, the number of these connections may vary. Better-known words are words with many connections, while poorly known words are words with few connections. In this way, any word which is encountered frequently in time develops a rich set of connections with other words.

Meara also points out that each individual acquisition event is small, but in the long term they add up to a rich vocabulary structure. When a word has limited connections to the main body of the lexicon, it is only part of the lexicon in a limited sense. It can affect the main lexicon in the sense that activation can spread from the word to the lexicon, but the word may not itself be affected by anything that happens in the lexicon. Meara explains that if we think of the links as connections that allow activity in one part

---

[12] Some signals are 'excitatory' (causing arousal), others are 'inhibitory' (causing suppression) (Aitchison, 1992:31).

of the lexicon to spread to other parts of the lexicon, then the pattern of connections enjoyed by the word means that it cannot share in the general activation patterns in the main lexicon (Meara, 1997:119). He also mentions that if the word could be activated by an external stimulus, then its activation could be passed on to the other parts of the lexicon.

Let us now look at some bilingual models of lexical processing and examine the way each model 'describes' L1 and L2 interaction. Green's (1986) initial model and its later version, the Inhibitory Control (IC) model[13] (Green, 1998), postulate that languages are organised as separate subsystems within the bilingual lexicon. These subsystems can have different levels of activation, allowing bilinguals to select the language they want to use. According to this model, to speak a language, this language has to be selected, (i.e. be the most highly active language) whereas the other language(s), if active, must be inhibited or suppressed. Individual lexical items bear language tags that correctly classify them in the appropriate language subsystem and activate or inhibit them accordingly.

De Bot's (1992) model is a bilingual version of Levelt's (1989) model of unilingual speech production where the processing of lexical-syntactic information takes place during grammatical encoding in the Formulator. Regarding the Formulator, then, De Bot assumes that it has a separate component for each language: in other words, grammatical and phonological encoding in L1 and L2 production entail different procedures. De Bot and Schreuder (1993:193) also see the necessity for an intermediate module ('Vbl') between the Conceptualizer and the Formulator "responsible for cutting up the fragment in chunks that can be matched with the semantic information associated with the different lemmas in the mental lexicon". Regarding the organisation of the bilingual lexicon, De Bot endorses Paradis's (1987) *subset hypothesis*, which assumes a single, language-independent lexicon, but in which L1 and L2 lexical items constitute different subsets and can be retrieved separately. Connections exist between lexical items of the same and of different (language) subsets. Different levels of activity may apply to the different subsets depending on the language being used at the time. De Bot and Schreuder (1993) also discuss code-switching data very much in a context of

[13] Green's model is the only one from those discussed in this section that is not formally based on Levelt's model.

exploring the issue of language separation in bilinguals. They refer to bilinguals' languages being turned on and off, and place in opposition Paradis's idea of separately activated subsets and Green's (1986) proposals of varying levels of activation.

Poulisse and Bongaerts' (1994) account of L2 speech production is again based on Levelt's model. They suggest that L1 and L2 lexical items are indistinguishably stored in a single, multilingual network but they bear a language tag signifying which language they belong to. Lexical items from both the same and different languages are interrelated within this network. The language tag is often the only distinguishing feature between translation equivalents from related languages. Contrary to de Bot's (1992) and Green's (1986, 1998) suggestion that lexical items are organised in different language subsets which are activated (or deactivated) as a whole, this model allows only a few lexical items from either L1 or L2 or both to be activated at a certain time. Lexical items are selected through spreading activation.

Considering the above, it is evident that there is as yet no integrated proposal about the nature of the bilingual lexicon. The existing accounts are primarily concerned with the distinction between different language systems or between lexical items from different languages. The organisation of the bilingual lexicon no longer concerns separate or common storage of L1 and L2 words, but at which levels and under what conditions L1 and L2 interact. The central issue about the nature of the bilingual lexicon refers to the particulars of the representation of a word at the conceptual/semantic and form levels; in other words, the mapping of meaning onto form.

1.5.3 Form and meaning in the L2 mental lexicon

One of the most important tasks of vocabulary acquisition in a L2 is the mapping of lexical forms to meanings. The question of the roles of form and meaning in the acquisition and processing of L2 lexis has been under recent debate. One claim is that the basis of the operations of the L2 lexicon is phonological rather than semantic, that "while in the native speaker's mental lexicon there are strong semantic links between the words, the connections between words in additional languages are primarily phonological" (Laufer, 1989:17 – referring to Fromkin, 1971; Hatch, 1983; Soudek, 1982). Singleton (1999) points out that the data which has been most frequently quoted

in support of the 'phonological' view of L2 lexical operations come from the Birkbeck Vocabulary Project word-association[14] tests, established and directed by Meara (see Meara, 1978, 1980, 1983, 1984). Meara provides evidence of associations made by native English speakers with a French stimulus word. All these associations illustrate some sort of phonological or orthographical confusion (Meara, 1984:233). Meara's argument is that the structure of the L2 mental lexicon is quite different from that of a native speaker's. The results of Meara's studies indicate that (a) the connections between words in the L2 learner's mental lexicon are less stable than the connections of the native speaker's, (b) phonology appears to play a much more prominent organizing role in the L2 mental lexicon than it does for native speakers[15], and (c) the semantic links between words tend to differ in a systematic way from those of native speakers (Wolter, 2001). In a similar pattern, Read (1993:358) summarises the results of word association studies as follows:

> *"One of the basic findings is that native speakers have remarkably stable patterns of word association, which can be taken to reflect the sophisticated lexical and semantic networks that they have developed through their acquisition of the language. On the other hand, second language learners produce associations that are much more diverse and unstable; often their responses are based on purely phonological, rather than semantic, links with the stimulus words. "*

Singleton (1999) points out that Meara's interpretation of his data can be criticised on the basis of a consideration of the nature of the Birkbeck tests. Although Meara (1984) claims that he presents his results as being founded on the use of 'very common' L2 items as stimuli, Singleton points out that some of the stimulus words are quite rare items such as *caque,* ('herring-barrel') and *émail* ('enamel') (Singleton, 1999:131). This suggests that the students did not actually display a L2 mental lexicon, which is qualitatively different from the L1 lexicon. It was a case of "a simple state of ignorance"

---

[14] Word associations are the links that connect or relate words in some manner in a person's mind. A common way of eliciting them is to have a tester give a prompt word and have the subject say the first word that comes to mind. Traditionally, researchers have been concerned with three types of responses on a word association test: paradigmatic, syntagmatic and phonological or 'clang' responses. 'Syntagmatic' associates are words which (frequently) collocate with the stimulus item (e.g. *dog-bite*, or *bark*) while 'paradigmatic' associates are words from the same word class as the stimulus item (the prompt word) having the same grammatical function within a given sentence. There are four main types of paradigmatic responses, including coordinates (i.e. the prompt word *dog* eliciting a response of *cat*), superordinates (*dog-animal*), subordinates (*dog-terrier*), and synonyms (*dog-canine*). 'Clang' responses resemble the prompt word only phonologically and bear no semantic connection to the prompt word, for example *dog-bog*).

[15] Meara (1983) noted that learners tend to produce 'clang' associations, nonrelated but similar-sounding words (*reflect-effect*), instead of the semantically related responses that adult native speakers typically produce.

(Singleton, 1999:132). In addition, Maréchal (1995) having re-analyzed data obtained by Meara (1978) found that only few of the responses reported by Meara are genuine responses related to stimuli in phonological terms only ('clang' responses). Further evidence on this issue comes from a recent study by O'Gorman (1996) of the English L2 word-association test responses of 22 Cantonese speakers. Although she was expecting to find evidence favouring Meara's view, she found from her data that the only clang associate to be found among her subjects' most common responses is *wealth* (in response to *health*). In all the other cases, the responses demonstrate clear semantic links with the relevant stimuli.

According to Singleton (1999), Meara's data can also be criticized on the basis of what is known about child/adult differences in performance on L1 word-association tests. Söderman (1989) places L2 data obtained by such tests in the context of L1 findings. Söderman argues that 'clang' associates are evident in children's L1 responses and implies that the proportion of phonologically motivated responses reflects the level of proficiency in a particular language rather than the type of this language in terms of nativeness or non-nativeness. Another attribute of child/adult L1 responses in word-association tests is a shift in response type distinguishable between children and adults. Söderman points to the lower proportion amongst adult responses of 'syntagmatic' associates and higher proportion of 'paradigmatic' associates. Native speakers (groups of older children) produced a higher proportion of paradigmatic responses. This shift in response type was related to some type of lexical or cognitive development. Thus, the phenomenon came to be referred to as either the syntagmatic-paradigmatic[16] shift or the developmental shift in response type (Wolter, 2001). Söderman (1989) claims that a shift in response type is also distinguishable between less proficient and more proficient stages of L2 development. She found that although native speakers did produce a higher proportion of paradigmatic responses than did the group of advanced learners for both lists (a high-frequency list and low-frequency list of adjectives of prompt words), in neither case were the differences statistically significant (Wolter, 2001). Additionally, both groups produced about an equal number of 'unusual' responses (which included 'clang' responses, and responses which could not be classified as either paradigmatic or

[16] A paradigmatic response is indicative of a higher degree of lexical or cognitive development than a syntagmatic response, which is indicative of a higher level of development than a clang or nonsensical response (Wolter, 2001:51).

syntagmatic) for words on the *high-frequency* list. Söderman, in a different experiment, found that the mean number of paradigmatic responses was positively related to proficiency (comparing non-native groups of ESL learners at different age and different stages in their experience of studying English). There was a shift in response type concerning the same English word-association test from proportionally more to proportionally fewer 'clang' associates and from proportionally more to proportionally fewer syntagmatic responses as L2 proficiency level increased. In brief, Söderman suggests that each lexical item passes from a more 'phonological' to a more 'semantic' profile as it becomes more integrated into the mental lexicon.

Singleton (1999) points out that the importance of meaning is also underlined by recent work on the place of phonological short-term memory in L2 lexical acquisition. Evans defines short memory as the capacity of the brain to hold information in a kind of immediate-access store for a short period after it has been presented (Evans, 1978:334). Short-term memory is used to store or hold information while it is being processed. It can hold information for only a matter of seconds. Another term to refer to the short-term phonological store is the term working memory. The object of vocabulary learning is to transfer the lexical information from the short-term memory to the more permanent long-term memory (long-term memory retains information for use in anything but the immediate future) (Schmitt, 2000). Singleton presents a number of studies showing evidence of the role of phonological representations in vocabulary learning, suggesting that the L2 mental lexicon differs from the L1 lexicon in being phonologically driven. For example, Ellis and Beaton (1993b), studying undergraduate English-speaking learners of psychology with no previous knowledge of German in attempts to memorize German lexical items, found significant correlations between the ease of pronunciation of foreign language words and their learnability.

Singleton (1999), however, points out that closer inspection of such evidence seems to suggest that these studies tend to confirm rather than challenge the view that the L2 mental lexicon resembles the L1 mental lexicon and that the 'phonological factor' in respect of vocabulary learning is prominent in the early stages of vocabulary learning in both L1 and L2. For example, there is evidence in L1 vocabulary learning, which indicates an important role for the short-term phonological store (see Gathercole and Baddeley, 1989, 1990, cited in Singleton, 1999:150). There is also evidence that the

semantic factor is present even in the very early stages of L2 vocabulary learning. Service (1993, 1993-94, cited in Singleton, 1999:150) suggests that L2 vocabulary learning depends on both phonological and semantic representations in working memory and on establishing connections between the representations in long-term memory.

Moreover, in the phonetic domain (of the mental lexicon), just as pre-verbal L1 learners have to struggle to replicate the sound shapes of their L1 (babbling), so L2 learners have to come to grips with sounds of the L2 that may bear little resemblance to those of their L1. In the conceptual/semantic domain, L2 learners, even if their exposure to the L2 begins in childhood, start from further down the road of concept development than infants confronting the task of L1 acquisition. This means that some of the concepts that have been lexicalized during L1 acquisition will be recyclable with only minimal adjustment in the L2 (Singleton, 1999:80) because the L2 learner already has experience of making relevant connections between lexical forms and meanings in his or her L1 (Singleton, 1999:48). In other words, the two major differences between the L1 and the L2 learner are that the latter, on the one hand, is at a more advanced stage of development in both physical and cognitive terms and, on the other, by definition, has already been through the process of acquiring a language (Singleton, 1999:79-80).

Considering the above, we reach the following basic conclusions (also cited in Singleton, 1999:167):

1) word-association test data fail to license a primarily 'phonological' conception of the L2 mental lexicon in contrast-distinction to a primarily 'semantic' conception of the L1 mental lexicon;
2) formal factors affecting L2 lexical acquisition also affect L1 lexical acquisition, and
3) in relation to the creation of L2 lexical memory codes, there is a meaning-focused dimension to even the earliest stages of this process.

In a similar pattern, Wolter (2001:45) draws the following conclusions and provides evidence for a structurally similar L1 and L2 mental lexicon:

1) Both native speakers of English and L2 learners demonstrate syntagmatic-paradigmatic shifts in responses.

2) Both native speakers of English learners of various levels of proficiency produce clang responses, mediated responses, and responses that seem completely unrelated to the prompt word.
3) A large diversity of responses can be found in the data of word association tests collected for L2 learners, NS (native speakers) adults, and NS children.

So, concerning the question of whether the L2 mental lexicon is intrinsically more form-based than the L1 mental lexicon, we notice that meaning rather than form poses the greater challenge. Lexical units (in both L1 and L2) are increasingly processed by meaning rather than form as their integration into the mental lexicon progresses. Regarding the question of whether or not there is connectivity between the L1 and L2 lexicon, it appears that L1 and L2 lexis are separately stored and that the two systems are in communication with each other – whether via direct connections between individual L1 and L2 lexical nodes, or via a common conceptual store (or both).

Due to the research presented above concerning word form and meaning in the L2 mental lexicon we could assume that in terms of development of lexical knowledge, learners move from mere word recognition, to partial and then to complete lexical knowledge. These last matters are fully presented in the next section.

1.6 Vocabulary development in a L2

In order to proceed with the discussion of vocabulary development in L2, it is important to clarify what 'knowing a word' actually means. The concept of 'knowing a word' has two dimensions (*lexical knowledge* and *lexical competence*) and two aspects (*breadth* and *depth*). It should be noted that these dimensions and aspects are interrelated in certain ways and degrees. This section examines this interrelation.

1.6.1 Lexical knowledge and lexical competence

The nature of lexical knowledge (that is the question of what it actually means for a language learner to 'know' a word) lies at the very heart of L2 vocabulary acquisition. The common distinction between knowledge and control (or competence and performance) is a possible source for terminological confusion. Meara (1978), for

example, describes this knowledge in behavioural terms as the ability to react to a word, while Henriksen (1999) argues for a competence-based description (see next section). We will now examine lexical knowledge and lexical competence according to Jiang's (2000) description.

As mentioned earlier, L1 words are learned as both semantic and formal entities but L2 words are learned as formal entities, because here the meaning is provided rather than learned from context. L2 learners' attention is focused on the formal features of the word. But, as Jiang (2000) points out, the suggestion that little semantic, syntactic and morphological information is represented in the lexical entry does not mean that these are not available to the learners. The meanings of L2 words and some grammatical information may become available through the activation of the L2-L1 link. Such semantic and grammatical information is not part of the mental lexicon. It is stored in the general memory or episodic memory and cannot be retrieved automatically. In other words, it is part of one's lexical knowledge, not one's lexical competence (Jiang, 2000). Jiang makes the distinction between lexical knowledge and lexical competence based on whether or not information is integrated into the lexical entry.

Lexical knowledge is "the knowledge or information a L2 learner remembers about the form, meaning, grammatical usage and sociolinguistic use of a word that is stored in a general memory system, rather than integrated into the lexical entry of a word" (Jiang, 2000:65). These different kinds of information are explicitly taught to and remembered by the learners in the process of learning a word. This information is available for conscious recollection.

Lexical competence, on the other hand, refers to "the semantic, syntactic, morphological and formal knowledge about a word that has become an integral part of a lexical entry in the mental lexicon and can be retrieved automatically in natural communication" (Jiang, 2000:65-66). For example, a L2 learner has lexical knowledge about a word but not lexical competence when he or she can state the rule regarding plurality but uses a singular form when a plural form is required. In this sense, lexical competence is not defined in terms of how much knowledge one knows about a word, but in terms of whether the knowledge is integrated into the lexical entry (Jiang, 2000).

1.6.2 Aspects of knowing a word

In the past decade or so, there have also been proposals that vocabulary knowledge could be regarded as having two primary aspects: breadth and depth (Qian, 1998, 1999; Read, 1988, 1989; Wesche and Paribakht, 1996). Breadth of vocabulary knowledge refers to the size of vocabulary or the number of words the meaning of which one has at least some superficial knowledge of. Depth of vocabulary knowledge relates to how well one knows a word. First, I will comment on breadth of vocabulary knowledge and then on depth, with special reference to Henriksen's (1999) three proposed dimensions for lexical competence.

1.6.2.1 Breadth of vocabulary knowledge

An important question for L2 vocabulary acquisition here is 'How many words does a L2 learner need?' The answer should be examined in relation to the number of words a native speaker knows and the number of words existing in the target language (Nation, 2001). Addressing the second question, the largest non-historical dictionary of English language is *Webster's Third New International Dictionary* which contains around 114,000 word families[17] excluding proper names. Nation (2001) mentions that this is a very large number of words and is beyond the goals of most first and second language learners. Regarding the number of words native speakers know, it has been suggested that educated native speakers of the English language know around 20,000 word families (see Goulden, Nation and Read, 1990; Zechmeister, Chronis, Cull, D'Anna and Healy, 1995). Addressing the main question, Nation (1990) proposes to focus on the 3,000 high frequency[18] words as an immediate priority. The classic list of high frequency words is the 2,000 word *General Service List* (West, 1953). These words are very important because they account for at least 85% of the words on any page of any book (Nation, 1990).

---

[17] A word family consists of a headword, its inflected forms and its closely related derived forms (Nation, 2001).

[18] *Frequency* refers to word frequency in a language. This topic is discussed later in the thesis (see section 2.3).

1.6.2.2 Depth of vocabulary knowledge

Depth of vocabulary knowledge will be examined as a part of Henriksen's (1999) three dimensions for lexical competence: (a) partial to precise knowledge, (b) depth of knowledge, and (c) receptive and productive knowledge. I will provide a description for each dimension:

(a) The **partial-precise knowledge** dimension describes different levels of comprehension of the same lexical item. On this continuum, vocabulary size is located toward the partial-knowledge end and more precise word knowledge would be found toward the precise-knowledge end. Word-recognition tasks have been used as very simple formats for measuring L2 vocabulary size. They only give an indication of whether or not a certain item is recognized as being part of the learner's vocabulary. Word recognition requires only the ability to recognize formal features of words; the learner may or may not reflect on meaning.

(b) The **depth of knowledge dimension** covers word knowledge components as found in other frameworks of vocabulary knowledge (e.g. Nation, 1990, 2001). Read (1993:357) defined the concept of depth in general terms as "the quality of the learner's vocabulary knowledge". Several studies have stressed the complexity of vocabulary knowledge and the many types of knowledge that comprise full understanding of a word (Schmitt, 1996; Wesche and Paribakht, 1996). Full mastery of a word requires more than just knowledge of its meaning and form (Nation, 1990).

Jiang (2000) points out that a great number of studies evaluated L2 vocabulary acquisition by measuring the percentage of new words the learners were able to recognize, recall, or provide definitions for (see Hulstijn, 1992; Ellis and Beaton, 1993a, 1993b; Griffin and Harley, 1996). These studies measured whether a word is remembered, rather than acquired. Schmitt (1998) suggests that one reason research has not yet translated into an adequate understanding of vocabulary acquisition is because nearly all of it has focused on the size (breadth) and growth of lexicons. Schmitt (1998) also points out that L2 vocabulary acquisition should focus on the acquisition of individual words rather than on the overall growth of the lexicon. To study the acquisition of individual words, one must be able to measure the degree or depth of

knowledge for each of the words. Schmitt provides two main approaches for doing this: 1) the developmental approach and 2) the dimension approach (Read, 1997).

1) The developmental approach uses scales to describe the stages of acquisition of a word. One scale that has received consideration lately is the Vocabulary Knowledge Scale (Paribakht and Wesche, 1993) which has 5 stages (see Table 1.1). Schmitt points out that although scales attempt to measure stages of knowledge in vocabulary acquisition, the definition of the stage boundaries may be problematic if acquisition is in fact a continuum. This is because it is difficult to give balanced attention to both receptive and productive knowledge using a scale. As Schmitt mentions, the beginning levels of a scale focus on receptive knowledge and the more advanced levels on production (Schmitt, 1998:285).

Table 1.1: Vocabulary Knowledge Scale (Paribakht & Wesche, 1993)

Stage 1: The word is not familiar at all.
Stage 2: The word is familiar but the meaning is not known.
Stage 3: A correct synonym or translation is given.
Stage 4: The word is used with semantic appropriateness in a sentence.
Stage 5: The word is used with semantic appropriateness and grammatical accuracy in the sentence.

2) The dimension approach describes the level of mastery of the various types of word knowledge. As already mentioned, the most complete and balanced description of word knowledge is that proposed by Nation[19] (1990:31). He developed a list of various types of knowledge that one must possess both receptively and productively in order to have complete command of a word. His description consists of eight word knowledge categories, each of which has receptive and productive aspects. Nation refers to:

1) The spoken form of a word;
2) The written form of the word;
3) The grammatical behaviour of the word;
4) The collocational behaviour of the word;
5) The frequency of the word;
6) The stylistic register constraints of the word;
7) The conceptual meaning of the word;
8) The associations the word has with other related words;

---

[19] Elaborating on Richards' list (1976).

Improving on his own framework (Nation, 1990), Nation (2001) proposed a set of 18 questions (about knowing a word) classified into three categories, each containing receptive and productive aspects: (a) form, including spoken form, written form, and word parts; (b) meaning, including form and meaning, concept and reference, and associations; (c) use, including grammatical functions, collocations, and constraints on use, such as register and frequency.

Similarly, Laufer (1997:141) mentions that it is generally agreed that knowledge of the following is necessary in order to know a word:

a. Form – spoken and written, that is pronunciation and spelling.

b. Word structure – the basic free morpheme (or bound root morpheme) and the common derivations of the word and its inflections.

c. Syntactic pattern of the word in a phrase and sentence.

d. Meaning: referential (including multiplicity of meaning and metaphorical extensions of meaning), effective (the connotation of the word), and pragmatic (the suitability of the word in a particular situation).

e. Lexical relations of the word with other words, such as synonymy, antonymy, hyponymy.

f. Common collocations.

Schmitt (1998, 2000) has commented on the strength and weakness of this type of list. He points out that concentrating on each of these word knowledge types for each individual target word would not be so practical in any normal classroom situation. But even though this approach is time-consuming and limits the number of words that could be studied, it would produce a very rich description of vocabulary knowledge (Schmitt, 1998). Such descriptive summaries can be used as frameworks for explanatory research, leading to a better understanding of how each lexical dimension is acquired (Schmitt and Meara, 1997).

As already mentioned, central to depth of knowledge is the process of network-building, which implies that learners create links between L2 words in their minds. The term 'network-building' is borrowed from L1 research. Aitchison (1994) described the acquisition of word meaning as a complex process, which involves: (a) labelling (or mapping according to Clark, 1993), (b) packaging, and (c) network-building. The first

term refers to the process of discovering which sequence of sounds can be used, as a name for a thing or entity (Henriksen, 1999:308). The second refers to the process of discovering the range of meaning for the same word and the third term refers to the process of discovering the sense relations[20] between words - that is, fitting the words together in semantic networks (reordering and changing of the lexical store by network building). There has been a tendency in L2 vocabulary acquisition research to focus on the mapping of meaning onto form and to disregard the aspect of network-building. Haastrup and Henriksen (2000), based on the idea that vocabulary acquisition is more a matter of system learning than of item learning, investigate the construct 'depth of knowledge' by providing the following phases leading to the integration of a word into a network.

1) **notice** phase: refers to the learner's ability to clarify meaning by detaching word meaning from context and focusing on its distinctive features
2) **analyse** phase: refers to the learner's ability to recognise and manipulate links between related L2 words in the lexical field. This includes both paradigmatic and syntagmatic relations
3) as the learner creates links between familiar L2 words and new L2 words, the new words gradually find their place within the lexical field. In order to **integrate** the new words in the existing network, the learner will have to continually structure and restructure relations between words.

According to Haastrup and Henriksen (2000), network-building is important for vocabulary acquisition. During this process of network-building, learners continually restructure their semantic network in order to accommodate new words. Haastrup and Henriksen (2000), point out that network-building is a slow process.

(c) The **receptive to productive knowledge**

The third dimension, according to Henriksen (1999), is receptive (R) to productive (P) use ability. The distinction between the receptive and productive vocabulary use needs a considerable discussion. The receptive/productive distinction resembles the distinction between the 'receptive' skills of listening and reading and the 'productive' skills of speaking and writing. Nation (2001) writes that receptive vocabulary use involves perceiving a word while listening or reading and retrieving its meaning. Productive

[20] For a detailed discussion on sense relations see Chapter Three.

vocabulary use, on the other hand, involves wanting to express a meaning through speaking or writing and retrieving and producing the appropriate spoken or written word form (Nation, 2001:24-25).

Clark (1993) and Meara (1990) argue for two dependent, but separate and qualitatively different R and P systems. On this view, R necessarily precedes P, the gap between R and P is large and a principled one, and R and P do not rely on identical information but are asymmetrical, different systems depending on different mental processes. Meara (1990) only accepts the notion of a continuum for P, while insisting that R is qualitatively different. He illustrates his point with a hypothetical graph structure (see Figure 1.8), where each word is represented as a node, and each association between words as an arc. Meara points out that in this example, node H, while clearly part of the overall network, is qualitatively different from all the other nodes in that it only has arcs pointing away from it, i.e. is inaccessible from anywhere else. Thus, H would correspond to an R item, which only responds to external stimuli, whereas P words can be activated by other words.

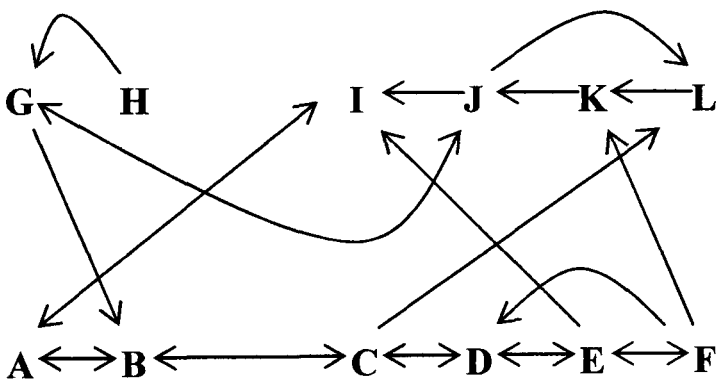


Figure 1.8 Meara's hypothetical association network.

In this sense, Meara's pattern suggests that the crucial distinction between *active* and *passive* vocabulary is that active vocabulary is connected to the lexicon by more than one type of connection. A word can become active as a result of a single exposure and it depends on which other parts of the lexicon are activated. In this way, if a word is connected to a part of the lexicon, which is not activated, then there is no activation to spread and the word will remain in a passive state. Corson (1995) also uses the terms active and passive vocabulary to refer to productive and receptive vocabulary and points out that the distinction is not based only on degrees of knowledge, because some

passive vocabulary may be very well-known but never used and therefore not active. Some people may be able to curse but never do (Nation, 2001).

Melka (1997:93), on the other hand, indicates that the distance between receptive (R) and productive (P) vocabulary is neither fixed nor permanent. She argues that there are three claims as far as the various sizes of the receptive and productive vocabulary are concerned.

1) R is much larger than P

There is evidence in L2 research showing that receptive vocabulary is larger than productive vocabulary (Marton, 1977; Clark, 1993). Erigna (1974) also found that the vocabulary of high school students after six years of French may be 4-5,000 words but their productive vocabulary may be 1,500-2,000 words.

2) the distance between R and P diminishes in the course of the learning career of the subject, though R remains larger

Morgan and Oberdeck (1930) conducted tests during the first semester of German learning at university level. Their study presents evidence that the distance between R and P diminishes slowly with the relatively slow development of productive knowledge.

3) the gap between R and P is non-significant: the two vocabularies are practically equal

At the end of their experiment, Morgan and Oberdeck found that though the gap between R and P was still evident, it was less significant than at the beginning of their experiment. Takala's (1984) results also suggest that Finnish learners of English have a receptive vocabulary not much wider than their active vocabulary.

On the basis of the previous studies, Melka (1997) suggests a new way of viewing the distance between R and P. She presents the distance as an image of numerous stages 'interrupting' the R and P line. She visualizes the distance between R and P as a line, a 'continuum of knowledge'. According to Melka, the idea of familiarity or degrees of knowledge is considered a measurement concept. The most elementary knowledge is the first encounter with a word in a context. At this stage, it is not clear if the word is stored in the mental lexicon at all. Higher degrees of knowledge which are close to productive knowledge include phonological, morphological, syntactical and lexical information about a word. But still it is not clear at which point receptive knowledge can be converted into productive knowledge. Melka argues that some aspects of the

word may have become productive, while others remain at the receptive level. Brown and McNeill's tip-of-the-tongue (TOT) experiment (1966) shows that subjects can comprehend[21] words even though they are stored in the lexicon in an incomplete way. Therefore, the passage from R to P is not so clear. Even when R is incomplete, P already begins. In this sense, the notions R and P overlap and R and P should not be considered as two separate systems functioning independently (Clark, 1993), but as one unique system used receptively or productively (Melka, 1997).

1.7 Depth of vocabulary knowledge and mental lexicon

Wolter (2001) claims that the L1 and L2 mental lexicons may in fact be similar, with depth of individual word knowledge determining a given word's degree of integration into the mental lexicon. He compares non-native and native speaker patterns of responses in light of depth of word knowledge scores and challenges the belief that a shift from predominantly syntagmatic to predominantly paradigmatic responses is indicative of lexical development. He attempts to devise and test a model of the L2 mental lexicon that is structurally similar to the L1 mental lexicon by looking at patterns of word associations of native and non-native speakers not only as a whole, but also with respect to how well the individual prompt words are known to the respondent.

Wolter (2001) suggests a Depth of Individual Word Knowledge Model. A DIWN model views the connections in both the L1 and L2 mental lexicons as conditioned not only by language proficiency, but by how well particular words are known to a given speaker (some words are known well, some not at all, and some are known to varying degrees). How well a particular word is known may condition the connections made between the particular word and the other words in the mental lexicon. Thus, it is possible that words in the lexicon are acquired individually and undergo developmental shifts separately from other words in the mental lexicon. In this sense, the mental lexicon can be viewed as consisting of a core vocabulary containing well-known words (including all that is entailed in 'knowing' a word, as described by Nation, 1990) and several layers of peripheral vocabulary consisting of words that are known to varying degrees (Wolter, 2001). In this model, the strength of connections between a particular word in the

---

[21] *Comprehension* can be defined in a narrow sense (decoding of linguistic information) and a broad sense (including extra-linguistic contextual information), see Clark et al., 1974, as cited in Melka, 1997:91.

mental lexicon and other words are conditioned by how well that particular word is known (in other word, its proximity to the core vocabulary). Thus, paradigmatic responses would be formed between words in the core, syntagmatic connections between words slightly further out, and phonological associations between words located on the peripheral layers (Wolter, 2001). For words that are moderately well-known, the syntagmatic links would be stronger than phonological ones, although both types would be present. Similarly, words in the core will have stronger paradigmatic connections than syntagmatic or phonological.

Wolters' findings suggest that paradigmatic responses do not in fact represent a higher degree of lexical development than syntagmatic responses. The progression for individual words could be seen as moving from a state in which phonological and other nonsemantic connections are dominant to a state where syntagmatic or paradigmatic connections take precedence (Wolter, 2001:65-66). According to the above model, phonology does play an important role in structuring non-native mental lexicon, for words that are moderately well-known. However, as words become well-known and better integrated into the mental lexicon, the phonological connections lose their predominance, and the semantic and syntagmatic links become stronger. In this sense, L2 mental lexicon is not less structured than the L1 mental lexicon, rather it is simply at an earlier stage of development.

1.8 Conclusion

It emerged from the above discussion that the mental lexicon is presented as a storage system which contains semantic and syntactic specifications (lemma) as well as morphological and formal (phonological and orthographic) specifications (lexeme) about lexical items. One important aspect of L1 lexical representation is that these different types of information automatically become accessible to native speakers. L1 learners put together the meanings of words by encountering them over time and in a variety of (meaningful) contexts. Successive encounters reinforce their grasp of what makes that word special, and help narrow down its range of operations in the language (Martin, 1984:130).

L2 lexical development and representation have also been discussed. Lexical development in L2 can be seen as comprising three stages. At the formal stage of lexical development a lexical entry is established in the L1 mental lexicon, but it contains only formal characteristics. During the second stage, a L2 entry becomes a combination of L2 formal information (in the lexeme) and the semantic and syntactic information of its L1 translation (in the lemma). At the final stage, semantic, syntactic, morphological, as well as formal specifications about a L2 word are established within the lexical entry. The fundamental difference between L1 and L2 lexical development is that the integration of semantic, syntactic, morphological and formal specification may occur for most, if not all, L1 words but only for a small proportion of L2 words. This happens due to limited contextualised input and the intervention of the existing semantic and lexical systems.

Other conclusions concern the role of form and meaning in the L2 mental lexicon. The word-association test data failed to license a primarily 'phonological' conception of the L2 mental lexicon in relation to a primarily 'semantic' conception of the L1 mental lexicon, and the formal factors which affect L2 lexical acquisition also affect L1 lexical acquisition (Singleton, 1999). Phonology does play an important role in structuring non-native mental lexicon, for words that are not well-known. However, the data suggest that as words become well-known and better integrated into the mental lexicon, the phonological connections lose their predominance and the syntagmatic links become dominant. Furthermore, as one's experience in the language increases, syntagmatic links lose their predominance and the paradigmatic links become stronger. Therefore, it is reasonable to think of that L2 mental lexicon as structurally different but not necessarily functionally inferior to the L1 mental lexicon.

In this section, vocabulary acquisition is viewed as a pattern which involves the building of a connection between a new word and a word that already exists in the learner's lexicon. Meara (1997) suggests that this connection might be a link between the new L2 word and its L1 translation equivalent, or it might be a link between the L2 word and an already known L2 word. On the basis of this assumption, Meara mentions that the crucial distinction between active and passive vocabulary is that active vocabulary is connected to the lexicon by more than one type of connection.

The dimensions (lexical knowledge and lexical competence) and aspects (breadth and depth) of knowing a word were also mentioned. We focused on depth by referring to the three dimensions for lexical competence proposed by Henriksen (1999): (a) partial to precise knowledge, (b) depth of knowledge, and (c) receptive and productive knowledge. There was a more detailed discussion about depth of word knowledge by pointing out that L2 vocabulary acquisition should focus on the acquisition of individual words rather than on the overall growth of the lexicon. To study the acquisition of individual words, one must be able to measure the degree or depth of knowledge for each of the words. Schmitt provides two main approaches for doing this: 1) the developmental approach, and 2) the dimension approach. According to Haastrup and Henriksen, network-building is important for vocabulary acquisition because it refers to the process of discovering the sense relations between words. They provide three phases leading to the integration of a word into a network: (a) notice phase (learner's ability to clarify word meaning), (b) analyse phase (learner's ability to recognise and manipulate links between related L2 words in the lexical field - this includes both paradigmatic and syntagmatic relations), and (c) integrate phase (learner's ability to create links between familiar L2 words and new L2 words, so that the new words gradually find their place within lexical field).

However, at this point, it is important to mention that contrary to the above views of breadth and depth of word knowledge as two separated notions, Vermeer (2001) argues that these two notions should not be considered two separated phenomena, but rather different dimensions of the same phenomenon. According to Vermeer (2001), there is the assumption that a deeper knowledge of words is the consequence of knowing more words. According to him, the ability to demarcate the precise meaning and usage of a word is based on knowledge of the other words that are needed to categorize, classify and define that word more precisely (Vermeer, 2001). In this way, depth of word knowledge (the ability to provide specific terminology, more associations, essential features, and functional characteristics about a word) is connected with breadth of vocabulary (knowing more words). He points out that breadth and depth should not be opposites but rather different dimensions of the same phenomenon, because if one knows more words, one can describe a stimulus word in greater depth. For example, if someone knows the words *cup*, *mug* and *glass*, he or she can describe the difference

between them and say that *cup* or *mug* is unlike *glass*. But if someone does not know the words *mug* and *glass,* he or she will call all three a *cup*.

In addition, Wolter (2001) suggests a Depth of Individual Word Knowledge Model which views the connections in both L1 and L2 mental lexicons as conditioned not only by language proficiency, but by how well particular words are known to a given speaker. How well a particular word is known may condition the connections made between the particular word and the other words in the mental lexicon. Wolters' findings suggest that paradigmatic responses do not in fact represent a higher degree of lexical development than syntagmatic responses. The progression for individual words could be seen as moving from a state in which phonological and other nonsemantic connections are dominant to a state where syntagmatic or paradigmatic connections take precedence. In this sense, the L2 mental lexicon is not less structured than the L1 mental lexicon, rather it is simply at an earlier stage of development.

Estimates of receptive (R) and productive (P) vocabulary were also discussed. It has been found that R is much larger than P, the distance between R and P diminishes in the course of the learning career of the subject, though R remains larger, and that the gap between R and P is non-significant: the two vocabularies are practically equal. Melka, (1997) presents the distance as an image of numerous stages 'interrupting' the R and P line. She visualizes the distance between R and P as a line, a 'continuum of knowledge'. R and P should not be considered as two separate systems functioning independently (Clark, 1993) but as one unique system used receptively or productively (Melka, 1997). To sum up, there is no complete, universally agreed upon description of R and P as yet, but it seems clear that it would have to account for the apparent lack of a neat divide between receptive and productive aspects of word knowledge.

After summarizing the main issues concerning mental lexicon and vocabulary development in L2, we can now proceed to the next chapter (Chapter Two) which provides a discussion on L2 vocabulary learning strategies and teaching methods.

CHAPTER TWO

Learning and Teaching Second Language (L2) Vocabulary

2.1 Introduction

Chapter Two is dedicated to vocabulary learning and teaching. The selection of vocabulary topics dealt with in this chapter corresponds to what I feel are the main vocabulary topics that are needed to outline the nature of L2 vocabulary learning and teaching. Current issues involved in L2 vocabulary learning and teaching will be discussed. This chapter begins with a discussion on different learning strategies and follows with an analysis of Nation's (2001) and Schmitt's (1997) taxonomy of vocabulary learning strategies. Then, the focus will be on the analysis of three different learning strategies that attract a great deal of attention from L2 researchers: guessing from context, dictionary look-up, and the keyword method. The importance of word lists in vocabulary teaching will also be discussed. The present chapter ends with an analysis of explicit vocabulary teaching methods with special attention to Sökmen (1997). The theory of the importance of meaning in L2 vocabulary development and the applications to vocabulary teaching will also be presented. Some of the methods and strategies outlined in this chapter are used in the current research as explained in detail in Chapter Five.

2.2 Vocabulary Learning

L2 language vocabulary acquisition is a very complex phenomenon involving several different learning processes. The most commonly drawn and pervasive distinction is that between *implicit* and *explicit learning*. Ellis (1994) has argued for a complete dissociation between the semantic and the formal aspects of vocabulary acquisition, claiming that the acquisition of semantic aspects of words (e.g. form-meaning connections) necessarily involves conscious, explicit learning, whereas the acquisition of the formal aspects of a word (e.g. phonetic and phonological features) is essentially implicit and unconscious in nature. This means that the semantic aspects of a word are more demanding of intellectual capacity, thus the formal aspects of learning a new word may be to the fore in the early stages but that after an initial concentration on form, the

longer-term task for the learner is to come to grips with a word's meanings and use (Singleton, 1999:153-54). Henning (1973, cited in Singleton, 1999), for example, found that L2 learners at a low level of proficiency registered vocabulary more by phonological similarities than by semantic relatedness, whereas high-proficiency learners relied on associated meanings rather than sound similarities.

The extreme positions regarding vocabulary acquisition processes, mentioned above, are expressed by the Implicit Vocabulary Learning Hypothesis and Explicit Vocabulary Learning Hypothesis, respectively. The former has its roots in Krashen's seminal *Input Hypothesis* (Krashen, 1989) and states that meanings of new words are acquired subconsciously as a result of repeated exposures in a range of contexts, where the conscious focus is not on form, but on message. According to the theory, learners acquire a L2 when they are exposed to comprehensible input which, in turn, is consistently effective in increasing proficiency. The latter holds that the employment of a range of vocabulary learning strategies (see next sections) can greatly facilitate and enhance vocabulary acquisition: on this view, learners are seen as active processors of information.

Krashen's Input Hypothesis has been extended in the form of the *Interaction Hypothesis* to explain how verbal interaction can create the conditions necessary for acquisition to take place (for a review see Ellis, 1995). *Input* refers to the linguistic forms used, *interaction* to the function served by those forms. When learners negotiate meaning by means of requests for clarification or confirmation checks, they can obtain interactionally modified input. This helps them to comprehend the input and focuses their attention on new learned linguistic forms, thus enabling their acquisition. Learners use and negotiate new vocabulary items in dialogically symmetrical discourse to create better conditions for vocabulary acquisition.

The notions of explicit and implicit learning should not be confused with the notions of incidental and intentional learning. *Incidental vocabulary acquisition* is learning vocabulary as the by-product of any activity not explicitly geared to vocabulary learning while *intentional vocabulary learning* is any activity aiming at committing lexical information to memory (Hulstijn, 2001). Although implicit learning can be incidental only (i.e. without learners' awareness of a forthcoming test, or without learners'

deliberate decision to commit information to memory), explicit learning can occur both intentionally and incidentally. Since linking 'word form' to 'word meaning' is an explicit learning activity requiring attention on the part of the learner, vocabulary can therefore be learnt intentionally as well as incidentally (Laufer and Hulstijn, 2001). Incidental learning occurs through extensive reading in input-rich environments (at a rather slow rate); however, the research makes it clear that this strong position is no longer tenable (Read, 2004:147). It is generally agreed that retention of new information depends on the amount and the quality of attention that individuals pay to various aspects of words. Hulstijn (2001:275), for example, argues that it "is the quality and frequency of the information processing activities (i.e. elaboration on aspects of a word's form and meaning, plus rehearsal) that determine retention of new information". Rich (qualitative) and numerous (quantitative) associations with existing knowledge (e.g. in the form of establishing similarities and contrasts between old and new information) increase the chances that the new information will be retained (Hulstiijn and Laufer, 2001). In practice, this means that if learners pay attention to the word's pronunciation, orthography, grammatical category, meaning and semantic relations to other words, they are more likely to retain the word than if they pay attention to only one or two of the above word properties. Thus, in the L2 classroom context incidental and intentional learning should be seen as complementary activities.

In addition, where vocabulary learning is more incidental to classroom activity, Laufer and Hulstijn (2001) argue that learning tasks[22] can be graded according to the level of vocabulary processing that they generate. They developed their Involvement *Load Hypothesis* for L2 vocabulary learning. The authors introduced a construct of involvement with motivational and cognitive dimensions. They propose that there are three factors in 'task-induced involvement':

- *need*: The learners' need to achieve a requirement. This is a motivational dimension of involvement. There are two degrees of prominence suggested for need: *moderate* (when it is imposed by an external agent, i.e. the teacher) and *strong* (self-imposed by the learners, for example, by the decision to look up a word in a L1-L2 dictionary when writing a composition),

[22] A task is defined here as "an activity or action which is carried out as the result of processing or understanding language (i.e. as a response)" (Richards, Platt, &Weber, 1985).

- *search*: Search for information on the meaning or form of the word (e.g. trying to find the L2 translation of a L1 word). Search is a cognitive dimension of involvement),
- and *evaluation*[23] : Evaluation deals with the way the information obtained applies to the particular use of the word in question, by comparing the word with other words in order to assess whether a word does or does not fit its context. For example, comparing homonyms, i.e. bank of a river, or bank as a financial institution. Evaluation is another cognitive dimension of involvement. From an analysis of previous research and their own parallel experiment (Hulstiijn and Laufer, 2001)[24], they found that tasks incorporating two or three of the factors led to better retention of the target vocabulary than those with only one factor. The combination of factors with their degrees of prominence constitutes involvement load. Hulstiijn and Laufer (2001) mention that teachers should design tasks varying in involvement load for different words depending on the type of reinforcement they want to provide.

2.2.1 What makes a word more or less learnable?

In order to enrich our understanding of vocabulary learning it is interesting to consider what makes a word more or less learnable.[25] Section 1.6.2 describes what is meant by knowing a word (e.g. word form or meaning). Knowing a word would ideally imply familiarity with all its features. However, in the case of vocabulary learning, knowing may be partial (i.e. the learner may have mastered some of the various types of word knowledge but not others). There are words which learners know in the sense of knowing what they mean in certain contexts, but which they cannot use productively. Other words vary in how easily they can be produced: some words can be retrieved only with effort; some are momentarily inaccessible (the tip-of-the tongue phenomenon).

---

[23] The kind of evaluation that entails recognizing differences between words is referred to as moderate. Evaluation that requires a decision as to how additional words will combine with the new word in an original sentence or text is referred to as strong evaluation (Hulstijn and Laufer, 2001).

[24] As they predicted, learners who wrote compositions using a set of target words remembered them better than those who encountered the words in a reading comprehension task and the learners who wrote the missing words in gaps in the reading text retained more of the words than those who just read marginal glosses.

[25] By learnable I mean that there are several features (properties which are related to the word's form and meaning) inherent in the word itself which might affect the ease or difficulty with which it is retainable or retrievable.

Laufer (1997:154) has reviewed a number of studies investigating a whole range of problems inherent in the nature of words themselves. She mentions several factors affecting word learnability. We will examine these factors one by one:

a) pronounceability (phonemes, combinations of phonemes, stress);

According to Laufer (1997), second language learners experience phonological difficulties related to phonemes, combinations of phonemes or stress. The L1 system may be responsible for learner's inability to discriminate between some phonemes because what makes a word phonologically more difficult than other is determined by the learner's L1 system. For example, Spanish speaking learners of English may have trouble distinguishing between pairs like *ban/van* or *day/they* while Hebrew speakers find it difficult to pronounce final consonant clusters in *clothes* and *films* (Laufer, 1997:143). Laufer (1997) reports a series of experiments (i.e. Gibson and Levin, 1975) which showed that pronounceable words are perceived more accurately than the unpronounceable ones.

b) orthography;

Orthography also affects word learnability because a different L1 writing system can be responsible for some learning problems. As Koda (cited in Coady and Huckin, 1997:44) points out, there are strong connections between the L1 orthographic system and L2 processing procedures. For example, native speakers of Semitic languages (which place great importance on consonants) tend to confuse words with similar consonants and different vowels (e.g. *pulls/pulse*) (Laufer, 1997:144).

c) length;

The misperception that longer words should be more difficult simply because there is more to learn and remembered is also reported by Laufer (1997). Coles (1982, cited in Laufer, 1997) found that long words produced more errors in recognition tasks than shorter ones. This suggests that shorter words are better learned than the longer ones. On the other hand, Laufer (1997) points out that if the components (morphemes) of the longer word are familiar (e.g. *unavailable*) there is no plausible reason why such a word should present a memorization difficulty. Another misperception is to assume that shorter words are easier because they are more frequent in the language. In English shorter words of Anglo-Saxon origin are indeed more frequently used that the longer words of Latin origin (Laufer, 1997:145). This does not happen in all languages. In Greek, for example, word length and frequency do not go hand-in-hand. The word *πραγματικότητα* (reality) is a long frequent word. It belongs to the first 1000 most

frequent words in the Greek language (Hellenic National Corpus, http://hnc.ilsp.gr/en/info.htm). In addition, Laufer (1997:145) points out that "even when length and frequency are related, what can account for better learnability is not the word's length, but the learner's frequent exposure to it".

d) morphology (inflectional and derivational complexity);

Morphology is another factor affecting word learnability. Features such as irregularity of plural, gender of inanimate nouns, and noun cases make a word more difficult to learn than a word with no such complexity. English speakers learning Hebrew, for example, find it difficult to master the Hebrew verb inflection (Laufer, 1997). Moreover, the lack of regularity with which morphemes can or cannot combine to create meaning can be another source of difficulty. For example, the learner must learn that *over* in *overthrow* can take a literal meaning but also *put an end to*; in *overcook,* it means too much (Laufer, 1997:146).

e) synformy;

Similar lexical forms are called 'synforms' (Laufer, 1997). There is evidence that L2 learners confuse words that sound and/or look alike. Meara (1982, cited in Laufer, 1997) found that some word associations of learners of French indicated that the stimulus word was confused with a similar-sounding word; i.e. the stimulus *béton* (concrete) elicited *animal*, which shows that *béton* was confused with *bête* (beast). Laufer (1985 and 1991, cited in Laufer, 1997) carried out a study of similar lexical forms and the extent to which they induced errors with foreign learners of English. She (1997) explains that general synformic similarity was defined in terms of the characteristics that all synforms shared (the identical number of syllables of the confused words, the identical syllabic position of the confused segments in the target word and error, identical stress patterns and part of speech); specific synformic similarities were classified into ten categories of synforms, each representing a different type of similarity between the target word and the error produced (i.e. category 1-synforms which have the same root, productive in present-day English but different suffixes, e.g. *considerable/considerate*). The ten categories are listed in Laufer (1997:147-148). Foreign learners of English were tested on all the possible synformic confusions and symformy proved to be an impeding factor for learners of English.

f) part of speech;

It is assumed that certain grammatical categories are easier to learn than others (i.e. nouns seem to be the easiest; adverbs, the most difficult; verbs and adjectives are

somewhere in between). Laufer (1997), referring to Odlin and Natalico's (1982) study, claims that there is no reported evidence for the prevalence of nouns or verbs over adjectives in terms of their learnability. Odlin and Natalico report lexical errors which show that, even though learners acquired the semantic content of some words, they confused their part of speech (i.e. nouns were replaced by verbs). On the other hand, Ellis and Beaton (1993b) claim that nouns are easier to learn as learners can form mental images of them more readily

g) abstractness;

It is often claimed that abstract words (e.g. *love*) are more difficult for L2 learners to learn than concrete words (e.g. *book*) because the former are more complex than the latter (see Allen and Vallette, 1972:114, cited in Laufer, 1997). On the contrary, Laufer (1997:150) argues that, while this is true in the case of L1 acquisition where lexical and cognitive development go hand in hand, L2 learners have already developed abstract concepts in their L1. For this reason, it cannot be claimed that concrete words are easier to learn because many abstract words may require simply learning a new form for a familiar concept (Laufer, 1997:150).

h) specificity/register restrictions;

Specificity and register restrictions may hinder word learnability. L2 learners do not often realize that lexical items frequent in one field of discourse may not be normal in another and that words acceptable when used with some interlocutors may not be suitable with others. This means that, words restricted to a specific register (co-hyponyms) are more problematic than general (superordinates) and neutral (e.g. in *vehicle/car*, *vehicle* is more general, *car* is more neutral) words, which can be used in a variety of contexts and registers (Laufer, 1997:151). Blum and Levenston (1978)[26] mention that learners preferred the Hebrew equivalent of *put* instead of *impose*.

i) idiomaticity;

Another factor that may affect the learnability of a word is idiomaticity. As Laufer (1997:151) admits, idiomatic expressions are much more difficult to understand and use than their non-idiomatic equivalents (*decide* is easier than *makes one's mind up*). This happens for two reasons; a) there is more that one word to learn and b) there is no clue as to the meaning of the idiom from the meaning of each word that builds it up (Laufer, 1997).

---

[26] In Laufer (1997:150).

j) multiplicity of meaning;

Multiplicity of meaning indicates that one word form can have several meanings (polyseme, e.g. *neck* or homonym, e.g. *bank*) and one meaning can be represented by different word forms. As Lyons (1981:148) points out, 'the problem of distinguishing between homonymy and polysemy is, in principle, insoluble'. Second language learners experience difficulties with synonymy and polysemy because, in practice, it is hard to distinguish which meanings are related and which are not (Laufer, 1997).

It is important to mention here that all the above factors are categorized as *intralexical* factors because they refer to intrinsic properties related to the word's form and meaning. Other factors that determine how difficult a word is for a learner are *interlexical* factors. These refer to the relationship between the L1 word and familiar words in the L2 (e.g. cognates[27]). Singleton and Little (1991) cite evidence indicating that the perceived degree of similarity between L1 and L2 will strongly influence the extent of transfer. Ringbom (1983) also notes that cognates can be extremely helpful to L2 learners, leading to positive transfer which affects reception in particular. The factor of word learnability as a part of vocabulary learning is to be seen in connection with vocabulary learning strategies which are discussed in the following sections.

2.2.2 What are language learning strategies?

Before looking at vocabulary learning strategies, possible characteristics of general language learning strategies will be discussed. There is no consensus on a definition of the term *language learning strategies*, probably because there are different interpretations of the terms *strategy* and *learning* in the literature. Wenden (1987:7) points out that strategies have been variously referred to as 'techniques', 'tactics', 'learning skills', 'potentially conscious plans', 'cognitive abilities', etc. Following Rubin's (1987:29) definition of learning as "the process by which information is obtained, stored, retrieved, and used", as well as her assumption that language learning is like other kinds of learning, we can say that *language learning strategies* could be any set of operations, steps or plans used by the learner which affect this process. Wenden (1987:6) identifies three areas in particular that language learning strategies

[27] Cognates are words with similarly formed translation equivalents in the L2.

refer to: (a) the actual behaviour of learners (what do learners do to learn a L2), (b) strategic knowledge (what do learners know about the strategies they use), and (c) knowledge about aspects (other than strategies) of the L2 learning process, such as personal/motivational factors. One of the first attempts at establishing a typology of language learning strategies was made by Rubin (1987). It was based on extensive empirical data collection in various settings. Another attempt, which provides a truly comprehensive overview complete with a hierarchical ordering of language learning strategies, can be found in Oxford (1990). Oxford (1990) organized the variety of vocabulary learning strategies into four strategy groups: Social Strategies (SOC) which promote interaction with other people to improve language learning, Memory Strategies (MEM) which relate new material to existing knowledge, Cognitive Strategies (COG) which exhibit the way the new language is used or manipulated or transformed by the reader, and Metacognitive Strategies (MET) which deal with conscious planning and evaluating the best way to study.

Language learning strategies are applicable to a wide variety of language learning tasks, ranging from vocabulary and pronunciation to grammar, speaking and reading comprehension. Individual learner differences are a crucial aspect in vocabulary learning strategies, as good learners vary in their choice of strategies and tend to use a wide variety of strategies in combination (Ridley and Singleton, 1995). As Schmitt and McCarthy (1997) point out, it is important to investigate how different learners combine different strategies and how this affects their learning and retention. In addition, it is worth examining if good strategies for vocabulary retention foster the development of general L2 proficiency and the way patterns of strategy users evolve and change over time as learners mature.

2.2.3 Nation's taxonomy of vocabulary learning strategies

Nation (2001) provides a taxonomy of three different kinds of vocabulary strategies. His taxonomy tries to separate aspects of vocabulary knowledge from sources of vocabulary knowledge and learning processes. The **first** general class of strategies (see Table 2.1) deals with planning vocabulary learning and involves deciding on where to focus attention, how to focus the attention and how often to give attention to the item (Nation, 2001:218). In this sense, students should be aware of their vocabulary goals and choose

what vocabulary to focus on according to these goals. They should have a clear strategy for deciding what vocabulary to focus on and where to find this vocabulary. They should also be aware of what is involved in knowing a word and be able to use a wide range of vocabulary learning strategies, bearing in mind that most vocabulary learning requires repeated attention to the vocabulary item.

Table 2.1: A taxonomy of kinds of vocabulary learning strategies (Nation, 2001: 218)

General class of strategies	**Types of strategies**
Planning: choosing what to focus on and when to focus on it	Choosing words Choosing the aspects of word knowledge Choosing strategies Planning repetition
Sources: finding information about words	Analyzing the word Using context Consulting a reference source in L1 and L2 Using parallels in L1 and L2
Processes: establishing knowledge	Noticing Retrieving Generating

The **second** class of vocabulary learning strategies involves finding information about words from a variety of sources. Nation points out that in order to learn new, unfamiliar vocabulary, learners have to be able to get information about the words. This information can be retrieved from the word form itself by analyzing its parts. Another way of finding information about words is to guess from context. Furthermore, consulting formal written sources (e.g. dictionaries or glossaries) and examining aspects of the relationship between L1 and L2 can also be useful for gaining information about words.

The **third** class of strategies involves ways of remembering vocabulary and making it available for use (Nation, 2001:221). Nation mentions three important general processes that may lead to a word being remembered:

- **Noticing**: It involves giving attention to an item. According to Nation, noticing requires decontextualisation. This means that the word is removed from its context to be focused on as a separate language item. Nation argues that this can occur in a variety of ways. Providing the definition of a word is one of them.

There are studies indicating that looking up words in a dictionary increases learning (Knight, 1994). Ellis (1995) suggests that simple definitions are the most effective, because they are short and they do not include many defining characteristics of the word. Nation suggests that the simplest kind of definition is L1 translation. There are strong arguments for using the learner's L1 if this will provide a clear, simple and brief explanation (Laufer and Shmueli, 1997). Richards and Taylor (1992) provide a list of definition types with examples showing the range of possibilities available (Table 2.2).

Table 2.2: Definition types

1)	Synonym	*beautiful means nice*
2)	Antonym	*young means not old*
3)	Analytic definition	An X is a Y which....
4)	Taxonomic definition	*autumn* is a *season*
5)	Definition by exemplification	*furniture* – something like a chair, sofa, etc.
6)	Definition by function	*pen* – use it to write
7)	Grammatical definition	*worse* – comparison form of bad
8)	Definition by association	*danger* – lives have not been protected
9)	Definition by classification	*family* – a group of people

- **Retrieval:** It involves recall of previously met items which may lead to a word being remembered (Baddeley, 1990). Nation argues that a word's specific form and meaning can be noticed and comprehended in a text through teacher explanation or dictionary use. Each retrieval of a word strengthens the path linking the written or spoken form of a word and its meaning or use. Retrieval may be receptive when the students perceive the form and have to retrieve its meaning when the word is met in listening or reading. It may also be productive when the students have to communicate the meaning of the word and retrieve its spoken or written form as in speaking or writing (Nation, 2001:67). At this point, Nation argues that there is a qualitative difference between studying words in notebooks (or lists) where the form, the meaning and the use of the word do not need to be retrieved, and retrieval strategies where students have to retrieve previously met information where only the word form is present and the other information has to be recalled by the students.
- **Generation**: It involves the process leading to a word being remembered. This process occurs when previously met words are met or used in different ways from the previous meeting with the word (Nation, 2001:68). Generation

> strategies include word analysis, semantic mapping, the use of grids and scales, collocations and sentences containing the word, mnemonic strategies (like the keyword technique), and using the word in new contexts of listening, speaking, reading and writing (Nation, 2001:222).

Nation points out that retrieving is superior to noticing because it encourages the establishment of vocabulary knowledge which is generation.

2.2.4 Schmitt's taxonomy of vocabulary learning strategies

Schmitt (1997) provides a very useful overview of the importance of L2 learning strategies. He attempts to present as complete a list of vocabulary learning strategies as possible. Schmitt's (1997) taxonomy is one of the most recent and comprehensive attempts to provide a classification scheme for a wide range of L2 vocabulary learning strategies, revising and expanding on Oxford's (1990) taxonomy in several important aspects: (a) it is especially geared to vocabulary learning and (b) compared to Oxford's typology the potential overlap of strategies is minimized. Schmitt's taxonomy is organized along two dimensions; first, Oxford's (1990) categories of general language learning strategies and second, the discovery/consolidation distinction.

The first classification dimension was adopted from Oxford (1990) who grouped learning strategies into four categories: (MET), (MEM), (COG) and, (SOC). Schmitt introduced a fifth category (DET). He argues that in Oxford's taxonomy there is no category including the strategies used by a learner when he or she discovers a new word's meaning without using any another person's expertise (Schmitt, 1997). He calls these strategies Determination Strategies (DET).

The second classification dimension was proposed by Nation (1990) and reflects the distinction of initial discovery of word meanings and remembering (as mentioned in the previous section). These strategies are divided into activities which are useful for a) the initial discovery of a word's meaning and b) remembering that word once it has been introduced. These strategies for gaining initial information about a new word are labelled Discovery Strategies and include DET strategies and SOC strategies (see Table 2.3). Schmitt suggests that once students have been introduced to a new item, it is

essential to use Consolidation strategies in order to remember it (these strategies could be SOC, MEM, COG or MET, as shown in Table 2.3).

Schmitt's taxonomy is listed in Table 2.3 and groups vocabulary learning strategies into 6 main categories with 58 individual strategies in total.

Table 2.3: A taxonomy of vocabulary learning strategies (Schmitt, 1997:207-208)

Strategy group	
Strategies for discovery of new word's meaning	
DET	Analyse part of speeches
DET	Analyse affixes and roots
DET	Check for L1 cognate
DET	Analyse any available pictures or gestures
DET	Guess from textual context
DET	Bilingual dictionary
DET	Monolingual dictionary
DET	Word lists
DET	Flash Cards
SOC	Ask teacher for a L1 translation
SOC	Ask teacher for paraphrase or synonym of new word
SOC	Ask teacher for a sentence including the new word
SOC	Ask classmates for meaning
SOC	Discover new meaning through work activity
Strategies for consolidating a word once it has been encountered	
SOC	Study and practise meaning in a group
SOC	Teacher checks students' flash cards or word lists for accuracy
SOC	Interact with native speakers
MEM	Study word in a pictorial representation of its meaning
MEM	Image word's meaning
MEM	Connect word to a personal experience
MEM	Associate the word with its coordinates
MEM	Connect the word to its synonyms and antonyms
MEM	Use semantic maps
MEM	Use 'scales' for gradable adjectives
MEM	Peg method
MEM	Loci method
MEM	Group words together to study them
MEM	Groups words together spatially on a page
MEM	Use new word in sentences
MEM	Group words together within a storyline
MEM	Study the spelling of a word
MEM	Study the sound of a word
MEM	Say new word aloud when studying
MEM	Image word form
MEM	Underline initial letter of the word
MEM	Configuration
MEM	Use keyword Method
MEM	Affixes and roots (remembering)
MEM	Part of speech (remembering)
MEM	Paraphrase the word's meaning
MEM	Use cognates in study
MEM	Learn the words of an idiom together
MEM	Use physical action when learning a word
MEM	Use semantic feature grids
COG	Verbal repetition
COG	Written repetition
COG	Word lists

COG	Flash cards
COG	Take notes in class
COG	Use the vocabulary section in your textbook
COG	Listen to tape of word lists
COG	Put English labels on physical objects
COG	Keep a vocabulary network
MET	Use English-language media (songs, movies, newscasts, etc.)
MET	Testing oneself with word tests
MET	Use space word practice
MET	Skip or pass new word
MET	Continue to study word over time

A number of interesting findings have arisen in Schmitt's (1997) research relating to his proposed taxonomy. Schmitt asked a total of 600 Japanese students and company employees with regard to actual strategy use and perceived helpfulness. The most popular vocabulary learning strategy proved to be consulting a bilingual dictionary; other well-liked vocabulary learning strategies were (in descending order) written and verbal repetition, studying the spelling of a word and guessing from context. On the other hand, strategies that were the least commonly used include using semantic maps, physical action, L1 cognates and the keyword method. Schmitt (1997) cautions, however, that these results may not be generalisable, as patterns of strategy use are likely to be culture-specific at least to some extent, and his subjects comprised Japanese learners only.

2.2.5 Analysis of three learning strategies

There are three learning strategies which I believe attract a great deal of discussion from L2 researchers: 1) guessing from context, 2) dictionary look-up and, 3) the keyword method. These strategies are further discussed below.

2.2.5.1 Results of studies of L2 learner's guessing from context

Schmitt (1997) argues that guessing an unknown word's meaning from context has been extensively promoted in the last two decades. Meara (1997) points out that everyone agrees that learners can acquire words by exposure to reading material, but everyone also agrees that the experimental evidence in support of this view is not strong because learners in these experiments acquire very few words (Hulstijn, 1992; Dupuy and

Krashen, 1993). These studies show that very limited gains are to be made from mere exposure to texts.

Sökmen (1997) presents various potential problems associated with inferring from context. She suggests that acquiring vocabulary mainly through guessing words in context is a very slow process, considering the fact that many L2 learners have a limited amount of time to learn this vocabulary. In this sense, maybe it is not the most efficient way to learn vocabulary (Carter and McCarthy, 1988; Scherfer, 1993). Haynes and Baker (1993) also argue that even when students use flexible reading strategies in order to guess words in context, their comprehension may still remain low because of insufficient vocabulary knowledge. Another problem is that guessing from context does not necessarily result in long-term retention. Parry's (1993) study of a university level ESL student's progress reading in English shows that the student was able to guess the correct meaning while working through a text but not when tested later. According to Coady and Huckin (1999:189), the main problems of incidental learning are; a) accurate guessing requires accurate word recognition, b) guessing takes time and thus slows down the reading process, c) guessing is effective only when the context is well understood (something that requires good textual clues) and d) guessing is not very effective in the acquisition of multiword lexical items.

Nation (2001) also mentions that guessing by non-native speakers has not shown large amounts of successful guessing and learning from context. In this case, there is a low rate of incidental vocabulary learning from context. Nation claims that this evidence has to be balanced against other considerations. We have to bear in mind that small gains become large gains if learners do large quantities of reading. Nagy (1997:27) claims that if a learner reads a million running words of text a year, and if two per cent of these words were unknown, this would amount to 20,000 unknown words per year. And if only one in twenty of these were learned, the students would still gain 1,000 words per year.

Furthermore, it has been argued that learning rates can be increased considerably by some deliberate attention to vocabulary (Elley, 1989; Hulstijn, 1992). This implies that direct vocabulary learning and incidental learning are complementary activities. For example, Fraser (1999) found that more vocabulary was retained from inferring from

context when the inferring was followed up by consulting a dictionary. Nist and Olejnik (1995) also found that when learners saw the word in context and then looked at a definition, the context helped their performance on a multiple-choice test. Furthermore, Paribakht and Wesche (1997) found that both the reading only and reading plus vocabulary instruction result in significant gains in vocabulary knowledge, but the supplementary vocabulary instruction produced more significant gains.

Extensive reading for meaning does not lead automatically to the acquisition of vocabulary. Guessing from context refers to guessing a word's meaning from the surrounding words in a written text and according to Schmitt (1997), it has two basic prerequisites, 1) richness of clues and, 2) level of language proficiency:

1) Richness of clues

It has been suggested that the context must be rich with clues that will help guessing (see Huckin, Haynes, and Coady, 1993; Nagy, 1997). Nation (2001) points out that learners have to be able to use the clues for guessing the unknown words. It is likely that at least 95% of the running words need to be already familiar to the learners for this to happen (Liu and Nation, 1985). This means that there is one unknown word in every 20 running words, or one in every two lines (Nation, 2001). According to him, to test the availability of context clues we need to focus on unknown words at the appropriate frequency level for the particular learners being tested. Furthermore, Bensoussan and Laufer's (1984) research indicates that some clues do not always help guessing and that many unfamiliar L2 words in a text for adult ESL readers have no contextual clues to meaning. It is essential that clues to the meaning of new words in the text should be clearly present, because the ease of learning these words depends on the clarity of their reference (Paribakht and Wesche, 1997).

2) Level of language proficiency

In order to guess from context, the learner must have a certain level of proficiency and be able to decode accurately the orthographical form of new words (Ryan, 1997) and have the knowledge of how to go through the inferencing process (Schmitt, 1997). According to Nagy (1997), three categories of knowledge contribute to context-based inferences: linguistic knowledge (e.g. syntactic knowledge), word knowledge and strategic knowledge (strategy instruction concerning the use of context). In the reading process there is a complex interaction between the reader and the text. Paribakht and Wesche (1997) point out that during this process the reader uses information from "the

surrounding text and from other knowledge sources to verify and elaborate the mental textual representation" (cited in Coady and Huckin, 1997:176). In this sense, they argue that the amount of cognitive processing required by the reader will be related to the 'depth' of comprehension of unknown words attempted by the reader. This will help establish the internalization of new knowledge about them, "with deeper processing leading to more acquisition" (Paribakht and Wesche (1997), cited in Coady and Huckin, 1997:176). So contextual guessing may be particularly helpful to students with higher proficiency.

2.2.5.2 Dictionary look-up

The second learning strategy to be examined here is dictionary look-up. Dictionaries can be used for a wide range of purposes. Nation (2001:281-82) provides the following list which covers most purposes for dictionary use:

(a) *Comprehension* (decoding): i.e. look up unknown words met, confirm the meanings of partly-known words and guesses from context.

(b) *Production* (encoding): i.e. look up the spelling, meaning, grammar, etc. of unknown or partly-known words needed to speak or write, correct an error, check that a word exists, etc.

(c) *Learning*: i.e. choosing unknown words to learn and enrich knowledge of partly-known words, including etymology.

Research findings are inconclusive as regards the benefit of using a dictionary for vocabulary learning. Knight (1994, cited in Nation, 2001:283), confirming Luppescu and Day's (1993) findings for a bilingual dictionary, showed that access to a (computerised) dictionary led to increased lexical knowledge after reading. In contrast, Bensoussan and Laufer (1984, cited in Zimmerman, 1997) demonstrated that the use of bilingual dictionaries does not significantly increase reading comprehension. Further, Hulstijn (1993) investigated the relationship between look-up behaviour and vocabulary knowledge on the one hand, and inference ability on the other hand. The result was that learners did not look up all unfamiliar words, but only those words that were most relevant to the reading comprehension task that they were set, ignoring words which were not relevant to the task. The ability to infer word meaning from context was not

related to look-up behaviour at all, which suggests that a substantial proportion of good guessers may have been inclined to check their guesses subsequently.

Moreover, Nesi (2000) claims that the amount of research on learners' preferences and their actual use of the dictionaries is remarkably limited. Nesi (2000) conducted a series of studies in which students looked up dictionaries as they performed either reading comprehension or sentence writing tasks. The results showed that the comprehension scores were not affected by whether the participants looked up words or not.

2.2.5.3 Keyword method

One of the most important mnemonic methods used in L2 teaching is the keyword method. When students find a word difficult to learn and remember, they are advised to use a mnemonic technique, which helps them to remember a new word's form with its meaning. According to Schmitt (1997), the keyword method is the most researched mnemonic strategy of all. It combines the phonological (or/and acoustic) forms and meanings of L1 and L2 words. The keyword method consists of three stages: (1) a L1 word is chosen based on acoustic and/or orthographic similarity with the L2 target word; (2) a strong association must be constructed between the target word and the keyword, so that the learner will remember immediately the keyword when he/she sees or hears the target word; (3) a visual image is constructed to combine the referents of the keyword and the target word (Hulstijn, 1997). For example, if an Indonesian learner wants to learn the English word *pin*, the learner could use the keyword *pintu* which is the Indonesian word for 'door'. The learner then thinks of an image involving a door and a pin (Nation, 2001). Kasper (1993:250) points out, that the students should concentrate on remembering the image between the keyword and the target word and that it is important to use different keywords for each target word

It has been found (Nation, 2001) that the technique works with learners of different achievement, learners at a variety of grade levels (including very young children), elderly learners and educationally disadvantaged learners. There are studies which show that the keyword method results in faster and more secure learning than other approaches (Moore and Surber, 1992; Brown and Perry, 1991). Some other studies have shown that the keyword method is also effective for recall of definitions, in sentence

completion tasks and in story comprehension (Avila and Sadoski, 1996). However, there is evidence which suggests that long-term retention is not good with the keyword method (Wang and Thomas, 1992 and 1995). Furthermore, Ellis and Beaton (1993a) mention that for successful productive recall and effective repetition the keyword needs to overlap considerably in form with the target word.

Hulstijn (1997) points out that the keyword method has a little appeal in the instruction of a foreign language, and it is rarely used (see also Kasper, 1993; Oxford & Crookall, 1990). There are textbooks where the keyword method is either not mentioned at all or only marginally (Nation, 1990; Taylor, 1990). Hulstijn provides three obvious reasons why the keyword method has not attracted too much attention. The first reason is that it can only be successfully applied with words referring to objects that can be perceived visually (concrete words). Although Kasper (1993) claim that the keyword method can be successfully applied with abstract words as well, Johnson (1985) found that the effectiveness of the method is significantly less when applied to abstract rather than concrete words. The second reason is that this method is less effective for the production of L2 words than for their reception, and the third most important reason is that teachers and textbook authors find this technique 'unnatural' or 'not serious enough' (Hulstijn, 1997).

2.3 Vocabulary Teaching: word lists and vocabulary size

Turning now to L2 vocabulary teaching, the L2 classroom teacher is faced with the challenge of how best to help students store and retrieve words in the target language. It is necessary to decide what vocabulary will be selected for teaching, how it will be sequenced and how it will presented. The first priority in (direct) vocabulary teaching is to focus on which words are to be studied. Vocabulary teaching is based on the type and number of words L2 learners need to learn. Nation (2001) mentions that learning a large number of words in L2 is not an essential short-term goal. While learning a very large number of words is useful in the long term, all words in a L2 are not of equal value to the learner. He points out that frequency-based studies show that some words are more useful than others. This is based on the type of vocabulary to be taught and learned. Vocabulary consists of four different types of words (Nation, 2001):

1) High-frequency words: These words include *function words* (*e.g. pronouns*) and *content words* (*e.g. nouns*). The classic list of high-frequency words is Michael West's (1953a) *A General Service List of English Words* which contains approximately 2,000 word families. These words are very important because they account for at least 85% of the words on any page of any book (Nation, 1990). For this reason, Nation suggests, teachers and students should pay attention and spend considerable time with them.

2) Academic and **3) technical words**: Academic words make up about 9% of the words in a text and contain words that are common in different kinds of academic texts (e.g. *sustain*). The *Academic Word List* (Coxhead, 1998) is very important for students using English for academic purposes. Technical vocabulary is very closely related to the subject area of the text. It is used within a specialized field (e.g. law, chemistry). These are words that are common in the specific topic area and not common elsewhere. Technical vocabulary is repeated within a text and is best taught within the content area of the relevant subject (Nation and Newton, 1997).

4) Low-frequency words: These words make up over 5% of the words in a text and have rare uses (e.g. *eponymous*). It is important to mention that, although it is not necessary to know low-frequency words to master a language, students should be encouraged to increase their vocabulary size by learning low-frequency words as well (Nation, 2001).

Apart from the different types of words, vocabulary teaching also depends on the size of vocabulary L2 learners need to know. Vocabulary size depends on deciding what words will be counted. Nation (2001) provides four different ways of word counting:

1. One way of deciding what words will be counted, is to count every word form in a spoken or written text, even if the same word occurs more that once. Words counted in this way are called '**tokens**'.
2. Another way of word counting is to count '**types**'. That is counting the words, but when we see the same word again in the text we do not count it.
3. The third way is to count '**lemmas**'. A lemma consists of a headword and some of its inflected and reduced forms (e.g. *n't*). English inflections include plural, third person singular present tense, past tense, past participle, *-ing*, comparative, superlative and possessive (Bauer and Nation, 1993, cited in Nation, 2001).

4. Another way of word counting is using '**word families**'. A word family consists of a headword, its inflected forms, and its closely related derived forms (Nation, 2001).

2.3.1 Explicit vocabulary teaching methods

According to Sökmen (1997), L2 research suggests that it is worthwhile to add explicit vocabulary instruction in the L2 classroom. In her attempt to examine how best to implement this kind of vocabulary, she considers the following pedagogical strategies:

1) *Build a large sight vocabulary*

As we mentioned earlier, L2 learners need to develop a large sight vocabulary, so that they may automatically access word meaning. As Sökmen (1997) suggests, whether they are self-selected words, difficult words, or high frequency words, the important thing for the students is to work consciously on the development of a large corpus of automatic word knowledge.

2) *Integrate new words with the old*

It has been suggested that the human lexicon is a network of associations, a web-like structure of interconnected links (Aitchison, 1987). In this sense, when L2 students connect the new word with already known words, the link is created and learning takes place. Teachers need to help L2 learners to establish these links and store vocabulary effectively. According to Sökmen, semantic mapping and charting semantic features help students to explore the relationship between the new word and words already known.

3) Provide a number of encounters with a word

Sökmen points out that if L2 learners meet the word in different contexts by using a variety of activities, a more accurate understanding of the word's meaning and use will be developed. A student needs a range of 5-16 encounters with a word in order to acquire it (Nation, 1990).

4) Promote a deep level of processing

Sökmen claims that one way to engage the L2 learners in deeper processing is to describe the target word until the meaning is clear. Another term used for this level of processing is 'rich instruction'. Nation (2001) argues that the aim of rich instruction is to establish the word as an accessible vocabulary item. This involves spending time on the word; explicitly exploring aspects of what is involved in knowing a word; and

involving the learners processing the word by group work or individual exercises (Nation, 2001:95). Nation points out that although many L1 researchers question the value of rich instruction, non-native speakers need vocabulary instruction because non-native beginners know very few English words. For this reason, it is practical to directly teach a number of high-frequency words. Nation provides two more reasons for using rich instruction in L2 vocabulary teaching. The first is that rich instruction increases L2 learners' opportunities to learn from rich vocabulary input, and the second is that direct, rich vocabulary instruction speeds up the learning process. This is very helpful for L2 learners since they have less time for learning. Nation provides several ways of providing rich instruction. Learners may examine a range of contexts containing the target word in order to provide a definition or translation. They may try to think of a L2 synonym or a L1 translation. They can also analyse the form and meaning of a word by breaking it into word parts and examine the etymology of the word. Furthermore, they can create lexical sets by putting the new and the known words into groups and grade them in some way. He provides (see Table 2.4) a range of activities for the various aspects of what is involved in knowing a word. Rich instruction involves giving attention to several of these aspects for the same word (Nation, 2001:97).

Table 2.4: A range of activities for vocabulary learning (Nation, 2001:99)

Goal		Activities
Form	spoken form	Pronounce the words Read aloud
	written form	Word and sentence dictation Finding spelling rules
	word parts	Filling word parts tables Cutting up complex words Building complex words Choosing a correct form
	form-meaning connection	Matching words and definitions Discussing the meanings of phrases Drawing and labelling pictures Peer teaching Riddles
	concept and reference	Finding common meanings Choosing the right meaning Semantic feature analysis

Meaning		Answering questions Word detectives
	associations	Finding substitutes Explaining connections Making word maps Classifying words Finding opposites Suggesting causes or effects Suggesting associations Finding examples
	grammar	Matching sentence halves Putting words in order to make sentences
	collocates	Matching collocates
Use		Finding collocates
	constraints on use	Identifying constraints Classifying constraints

5) *Facilitate imaging and concreteness*

Sökmen refers to the dual theory of human memory (Clark and Paivio, 1991) and points out that the mind contains a network of verbal and imaginal representations for words. When learners image the new word, the possibility for later recall is much greater than if they only make verbal links. Some ways to build non-verbal representations are: make illustrations, show pictures, draw diagrams and list details (Sökmen, 1997:244).

6) *Use a variety of techniques*

Sökmen provides a great number of instructional techniques and suggests that teachers should use a mixed approach with these techniques because it breaks up the class routine while building a variety of associational links. Sökmen divides these strategies into six categories. I will refer to these categories but not in detail:

i) 'Dictionary work' is focused on the word and its definition. She provides good examples of 'dictionary work' which help students to expose themselves to a variety of ways to practise words and their definitions (see Sökmen, 1997:245).

ii) Word unit analysis. Attention to word parts allows learners to make use of the word families they know and remember new complex words (Nation, 2001).

iii) Mnemonic devices. These devices are efficient in storing words. Sökmen argues that the most common verbal mnemonic device is using the rhyming of poetry or song.

The most often studied mnemonic device is the keyword technique, which employs both an acoustic and a visual image.

iv) Semantic elaboration

a) Semantic feature analysis

Channell (1981) argues that analysing the meaning components of words is very effective in teaching and language. Nation (2001) claims that numerous writers suggest that learners should fill in grids in order to refine their knowledge of related words. This kind of activity can be used when learners are familiar with most of the items being compared (Nation, 2001:103).

b) Semantic mapping

With this activity, it is possible to analyse words in different ways because it refers to brainstorming word association and then diagramming the results.

c) Ordering

When students arrange a list of words in a specific order, they integrate new information with old, establishing memory links

d) Pictorial schemata

Another semantic strategy is to create grids or diagrams to encourage lexical ordering. The teacher can use scales, Venn diagrams or tree diagrams.

Nation (1990) points out, that semantic techniques are better used as review techniques.

v) Collocations and lexical phrases

Practising collocations (words that commonly go together) is a worthwhile activity because they have very powerful links in the lexicon (Aitchison, 1987). Nattinger and Decarrico (1992) define these phrases as 'ritualized bits of language'. Learning them as chunks is more efficient than learning individual words. They suggest starting with a small number of phrases, practising them in pattern drills and then in single, predictable situations.

vi) Oral production

Oral activities using the new vocabulary break up the class routine and the students experience words with oral reinforcement. According to Nation (1990), dialogues have the advantage of putting words directly into productive vocabulary. Role-playing, which

is a less practised technique, is an opportunity for more spontaneous oral practice of vocabulary (Sökmen, 1997).

2.4 Conclusion

It emerges from the above discussion that L2 vocabulary learning is a complex phenomenon and involves different types of learning such as explicit/implicit and incidental/intentional. Furthermore, extensive work has been cited regarding different taxonomies of learning strategies. The first solid attempt was made by Oxford (1990) who organized strategies around Social, Memory, Cognitive and Metacognitive Categories. A different approach was provided by Nation who provides a taxonomy of three different kinds of vocabulary strategies: the first general class of strategies deals with planning vocabulary learning and involves deciding on where to focus attention, how to focus the attention and how often to give attention to the item (Nation, 2001:218). The second class of vocabulary learning strategies involves finding information about words from a variety of sources. The third class of strategies involves three ways of remembering vocabulary and making it available for use: 'noticing' (giving attention to an item), 'retrieval' (recall of previously met items) and, 'generation' (the process leading to a word being remembered). As a combination of the two previous classifications concerning learning strategies, Schmitt (1997) developed an extensive taxonomy of learning strategies. He added the DET strategy to the ones existing in Oxford's and used Nation's taxonomy to further classify learning strategies. Schmitt's taxonomy consists of 58 strategies divided into 6 main categories.

There was also a brief discussion on three learning strategies which I believe attract a great deal of discussion from L2 researchers: 1) guessing from context, 2) dictionary look-up and, 3) the keyword method. As far as the results in studies of L2 learners' guessing from context is concerned, Sökmen (1997) presents various potential problems. She mentions that acquiring vocabulary mainly through guessing words in context is a very slow process, considering the fact that many L2 learners have a limited amount of time to learn this vocabulary. It has also been suggested that for guessing the meaning of unfamiliar words in context, clues are needed. Nation (2001) points out that L2 learners have to be at certain level of proficiency in order to use the clues for guessing the unknown words. The findings from studies on the effectiveness of using a

dictionary and the keyword method are not conclusive. Concerning dictionary look-up, research findings are inconclusive as regards the benefit of using a dictionary for vocabulary learning. Moreover, Nesi (2000) claims that the amount of research on learners' preferences and their actual use of the dictionaries is remarkably limited. It has also been found that there are studies which show that the keyword method results in faster and more secure learning than other approaches. However, there is evidence which suggests that long-term retention is not good with the keyword method (Wang and Thomas, 1992, 1995). In addition, Hulstijn (1997) points out that the keyword method has little appeal in the teaching of a foreign language, and it is rarely used.

It was also stated that vocabulary teaching should be based on the type and number of words L2 learners need to learn. Nation (2001) mentions that some words are more useful than others. This is based on the type of vocabulary to be taught and learned. Vocabulary consists of four different types of words: High-frequency words, Academic words, Technical words and Low-frequency words. Apart from the different types of words, vocabulary teaching also depends on the size of vocabulary L2 learners need to know. Vocabulary size depends on deciding what words will be counted. Nation (2001) provides four different ways of word counting: counting words as tokens, types, lemmas and word families.

According to Sökmen (1997), L2 research suggests that it is worthwhile to add explicit vocabulary instruction in the L2 classroom through a series of pedagogical themes which deal with: how to integrate new words with old words; how to promote a deep level of processing using rich instruction, and how to use a variety of techniques. These kinds of teaching methods direct learners' attention to lexical relations and support the theory of the importance of meaning, depth of knowledge and network-building in L2 vocabulary development. We notice that in learning strategies and teaching methods sense relations (i.e. synonyms, antonyms) are used extensively in order to promote a deep level of word processing. Semantic mapping, for example, helps students to refine their knowledge of related words. The process of discovering the sense relations (network-building) between words is central to the depth of vocabulary knowledge (as mentioned in section 1.6.2.2). Some of the sense relations are presented in detail in the following chapter (Chapter Three).

CHAPTER THREE

Sense Relations

3.1 Introduction

The third chapter will be looking at lexical semantics and more specifically at sense relations mainly for two reasons. First because, as mentioned in previous chapters, sense relations facilitate depth of word processing and knowledge (as appear in both learning strategies and teaching methods) and, second because specific relations will be used in the teaching methodology of the present study (described in Chapter Five). Semantics is the technical term used to refer to the study of meaning (Lyons, 1995) and the unifying theme of lexical semantics is the idea that we can state the meaning of words in terms of their associations with other words (Palmer, 1981).

The main purpose of the chapter is to discuss the following six kinds of sense relations between words: polysemy, homonymy, hyponymy, antonymy, meronymy and synonymy. My main objective is to focus on sense relations between words not between pairs of sentences. The focal point of this chapter will be the analysis of synonymy as a sense relation between words. The analysis of synonymy cannot stand on its own. It is important to examine synonymy by considering other kinds of sense relations as well. The discussion of the other five sense relations helps to show how such relations are essential in language knowledge and helpful for explaining the meaning of words. Understanding these relations is useful for introducing new vocabulary and for creating activities to enrich L2 learners' understanding of words (as mentioned in Chapter Two).

The relationships described here are useful starting points for devising classification activities with words that learners already know. Classification activities can also involve distinguishing and grouping similar items in various ways, justifying the distinguishing and grouping by explaining the relationships. In other words, the presentation and teaching of synonymous, antonymous or homonymous terms depends on the explanation of the relationship between these terms. The second chapter will begin with an examination of what is meant by the term *word* and sense relations in general and then it will investigate the six sense relations mentioned above.

3.2 'Word' definitions and sense relations

Lyons (1995:80) defines the sense of an expression as the set or network of sense-relations that hold between that expression and other expressions of the same language. At this point, it is important to clarify the term *expression* that Lyons uses by examining different definitions of the term *word.*

According to Singleton (1999:10), what is meant by the term *word* will depend on the level of abstraction at which a given speaker or writer is operating, the linguistic 'level(s)' and the extent to which semantic content is being treated as criterial. With regard to level of abstraction, there are two points. On the one hand, words can be thought of in terms of types, tokens, lemmas, and word families (explained in section 2.3). On the other hand, there is a usage of *word* according to which the phrase *Going, going, gone* would be judged to contain just one word- the verb *go*, represented by two of its various forms (*going* and *gone*). The abstract unit based on a collection of forms thus seen as constituting in some sense a singular lexical entry is often referred to as a *lexeme* or a *word expression*, while its concrete 'representatives' are referred to as *word forms.* Thus BRING is the lexeme which underlies different grammatical variants: 'bring', 'brought' which can be referred to as word-forms. When we look up words in a dictionary we are looking up lexemes rather than words. That is, 'brought' and 'bringing' will be found under an entry for BRING. For practical purposes, lexemes (dictionary entries) have conventional citation forms; i.e. an English noun lexeme will be cited by the use of a singular form (*woman, sea*).

As far as words and linguistic 'levels' is concerned, there are several approaches to illustrate what is meant by the term *word.* Singleton (1999) provides the following example. Consider the word *builds*. One approach is to identify the word in orthographic terms – as a sequence of letters bounded on either side by a blank space or punctuation mark - b + u + i + l + d + s builds. A second approach is to identify the word as a phonetic entity with its particular acoustic characteristics having to do with the modalities of its production. Another way is to present the word as a phonological entity, as a sequence of units which are functionally relevant in the sound structure of English - /b/ + /I/ + /l/ + /d/ + /z/. At a morphosyntactic level it is the third person

singular form of a verb; and at a semantic level it is (among other things) synonymous with *constructs*.

Each lexeme is related, in various ways, to other lexemes and expressions of the same language. The lexeme 'dog' for example, is related, in various ways, to other lexemes, like 'animal', 'hound', 'terrier', etc. and each such relation that holds between 'dog' and other expressions of the same language system, is identified as one of its sense-relations (Lyons, 1995:79). Sense-relations are divided into substitutional (paradigmatic: relations between members of the same grammatical category) and combinatorial (syntagmatic: between expressions of different grammatical categories i.e. between nouns and adjectives) which can be put together in grammatically well-formed combinations (Lyons, 1995:124).

3.2.1 Polysemy and homonymy

Hurford and Heasley (1983) identify polysemy when the same word (word form) has several very closely related senses. *Mouth* (of a river vs of an animal) is a case of polysemy. Homonyms are word forms pronounced or spelt in the same way but having unrelated senses far apart from each other and not obviously related to each other in any way. *Mug* (drinking vessel vs gullible person) is a clear case of homonymy. Saeed (1997) points out that some authors divide homonyms into homographs, word forms with identical spelling: *lead* (metal), *lead* (dog's lead) and homophones, word forms pronounced the same way: *by*, *buy*.

Palmer (1981) identifies a practical problem concerning polysemy. When one form has several meanings, it is not always clear if this is a polysemic word or if this is a case of homonymy. Palmer provides the following example: while the dictionary defines the word *flight* as a single, polysemic word ('passing through the air', 'power of flying', 'air journey', etc.), it provides five word forms (homonyms) for *mail* – 'armour', 'post', 'halfpenny', 'payment' and 'spot'. These words are not shown as different meanings of the same word. Palmer points out that, the dictionary has to decide whether a particular item is to be handled as a case of polysemy or homonymy, because if it is a case of polysemy a single entry will be provided, and if it is one of homonymy, a separate entry has to be provided for each of the homonyms (Palmer, 1981:101). The notion of lexeme

helps us to represent polysemy in individual words: thus, *fair* (n.), *fair* (adj.1 as in good, acceptable) and *fair* (adj.2 as in light in colour, especially of hair), would have three different lexeme meanings for the same word form.

The problem is to decide when we have polysemy and when we have homonymy. Saeed (1997) points out that, lexicographers tend to use criteria of 'relatedness' to identify polysemy. Speakers' intuitions and the etymology (what is known about the historical development of a word) are included in these criteria. Sometimes, the etymological criterion supports the native speaker's intuitions about particular words (or lexemes). For example, as Lyons (1995) point out, native speakers would classify *bat* ("furry mammal with membranous wings") and *bat* ("implement for striking a ball in certain games") as two different lexemes. These words do differ in respect of their etymology. *Bat*1 is derived from a regional variant of Middle English 'bakke', and *bat*2 from Old English 'bat' meaning "club, cudgel" (Lyons, 1995:59). But etymology does not always support the intuitions of native speakers. Saeed (1997) implies that speakers' intuitions may conflict with the historical facts. He provides the following example with the words *sole* 'bottom of the foot' and *sole* 'flatfish' (Saeed, 1997:65). Most English speakers feel that the two words are unrelated and that they should appear in the dictionary as homonymous, providing different lexical entries. But these words are actually derived (via French) from the same Latin word *solea* 'sandal'. We can argue on this ground that we have a case of polysemy. In this case, dictionaries follow speakers' intuitions and list these words separately.

Lyons (1995) distinguishes absolute homonymy from partial homonymy. According to him, absolute homonyms satisfy the following three conditions:

(i) they will be unrelated in meaning;

(ii) all their forms will be identical;

(iii) the identical forms will be grammatically equivalent.

Absolute homonyms are *bank* ('financial institution') *bank* ('sloping side of a river'), *sole* 'bottom of the foot' and *sole* 'flatfish'. Lyons also identifies partial homonymy where a) there is identity of one form and b) one or two, but not all three of the above conditions are satisfied (Lyons, 1995:55). He presents the following example. The verb *find* and *found* share the form 'found', but not 'finds', 'finding', or 'founds', 'finding',

etc.; and *found* as a form of 'find' is not grammatically equivalent to *found* as a form of 'found'.

3.2.2 Hyponymy

Palmer (1981:85) points out that hyponymy refers to the notion of 'inclusion'. For example *tulip* and *rose* are included in *flower*, *scarlet* is included in *red*. Palmer sees inclusion as a matter of class membership, where the 'upper' term (*flower*) is superordinate (or hypernym, as it is termed in Saeed, 1997:68) and the 'lower' term (*tulip*) the hyponym. Hyponymy is a relationship existing between specific and general lexical items. The meaning of the specific item is included in, and by, the meaning of the more general item. *Tulip* and *rose* are co-hyponyms and are linked by their common inclusion under a superordinate (or hypernym) *flower* in whose class they belong. The sense of *tulip* includes that of *flower*. In other words, *tulip* entails *flower*. Entailment is a relation that holds between propositions. A proposisition is that part of the meaning of the utterance (real piece of speech) of a sentence (abstract grammatical elements obtained from utterances) which describes some states of affairs (Saeed, 1997). Lyons (1995) suggests that it is convenient to be able to say that one word or phrase entails another, just as it may be convenient to be able to say that one sentence entails another. Based on this, we can say that X, is a hyponym of Y if and only if X entails ($\Rightarrow$) Y, i.e, *dog* $\Rightarrow$ *animal*. Hyponymy is better illustrated in the following example borrowed from Saeed (1997:69). We see that *hacksaw* is a hyponym of *saw* and *saw* is a hyponym of *tool* (see Figure 3.1). We also notice that *tool* is a hypernym of *hammer*, *saw* and *chisel* which are co-hyponyms.

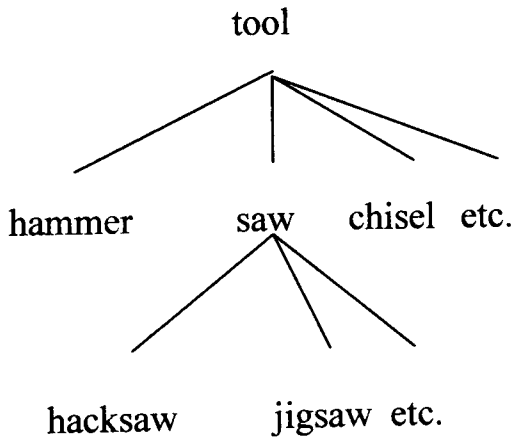


Figure 3.1 Example of hyponymy (Saeed, 1997:69).

3.2.3 Antonymy

The traditional terminology suggests that antonyms are words which are opposite in meaning. The term antonymy is used for 'oppositeness of meaning' (Palmer, 1981) and it is exemplified by such pairs as *long/short*, *good/bad*, *hot/cold* (Cruse, 1986). The scale on which a pair of antonyms operates is not a simple one, however. In order to show how a pair of antonyms works, Cruse (1986) refers to two scales: an absolute scale, which covers all possible values of the scaled property from zero to infinity, and a relative scale, which is movable relative to the absolute scale, and whose values are directly relatable to the terms of the antonymous pair. In this way, one term will tend towards zero and the other will tend in the contrary direction. Cruse presents the way a pair of antonyms operates with the following diagram (see Figure 3.2). It refers to the antonyms *slow/fast*.

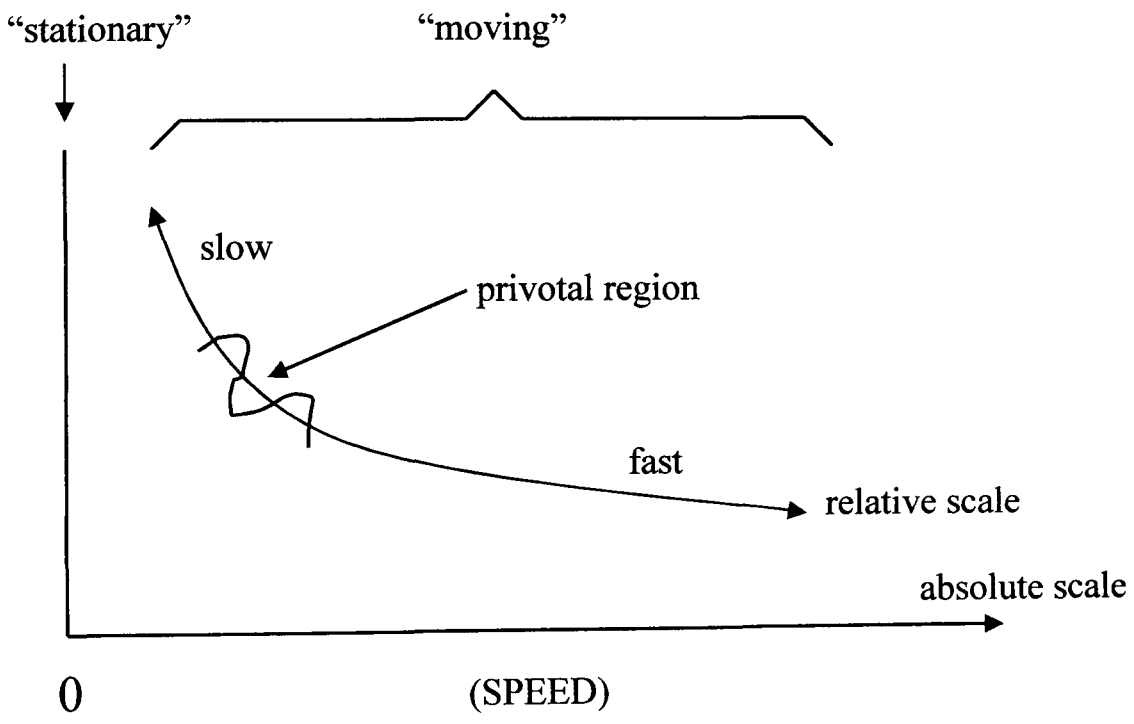


Figure 3.2 The scale on which a pair of antonyms operates (Cruse, 1986:205).

The pivotal region is the neutral region of the scale, where the terms of antonymous pairs are symmetrically disposed. This region cannot be referred to by either member of the pair, and in most cases it is not designated linguistically by any lexical item. Cruse (1986:205) points out that *tepid* and *lukewarm*, referring to the pivotal region between *hot* and *cold*, are exceptional.

Cruse explains that the vertical dimension in the diagram is not important; it simply permits the representation of an important property of *slow*. The value of *slow* tends towards zero speed but never actually reaches it, because this is not a physical fact, "but a linguistic one: we cannot say *completely slow* when we mean 'stationary'" (Cruse, 1986:206). For this reason we cannot say *completely short* when we mean 'having zero length'.

Saeed (1997) sees antonymy under a more general label of opposition. He provides the following (five) categories for antonyms (see Table 3.1):

1. The first category includes simple antonyms, which are also called complementary pairs or binary pairs. The positive of one term implies the negative of the other – e.g. *dead/alive*, *pass/fail*. So, dead implies not alive.

Table 3.1: Different categories (senses) of antonymy

Types	Examples
complementary pairs or binary pairs	*dead/alive*
gradable antonyms	*hot/cold*
reverses	*push/pull*
converses	*own/belong to*
taxonomic sisters	*Monday/Sunday*

2. The second category is called gradable antonyms. In these opposite pairs, the positive of one term does not necessarily imply the negative of the other i.e. *rich/poor*, *fast/slow*. He points out that gradable antonyms mainly include adjectives and have two major characteristics: a) there are usually intermediate terms i.e. between the gradable antonyms *hot* and *cold* we can find the words *warm*, *tepid*, *cool* and b) the terms are usually relative i.e. *a thick pencil* is likely to be thinner than *a thin girl* (Saeed, 1997:67). A good test for gradability, is to see if the word can combine with *very*, or *very much*, or *how*? or *how much*? (Hurford and Heasley, 1983).

Palmer (1981) provides some characteristics of some gradable antonyms and their dichotomy. There are some antonyms, e.g. *honest/dishonest, open/shut* that are gradable in terms of *more* or *less*. With these antonyms the denial of one frequently is taken to assert the other. Palmer provides the following example (1981:97). Though, we may say (1) 'Bill is more *honest* than John', (2) 'Bill isn't *honest*' implies that (3) 'Bill is *dishonest*', and (4) 'Bill isn't *dishonest* implies' that (5) 'Bill is *honest*'. These antonyms are explicitly gradable, but as Palmer says, they are not usually treated as implicitly gradable. Another characteristic is that the *more* and *less* relationship cannot be applied to some pairs of antonyms. For example, the antonyms *brilliant/stupid. More brilliant* does not equal *less stupid.* As Palmer points out, these terms, although gradable, also have an absolute value at one of the 'ends' of the scale.

3. The next category that Saeed (1997) provides includes terms called reverses. The characteristic reverse relation is between two opposite terms describing movement, where one term describes movement in one direction, and the other the same movement in the opposite direction – e.g. *push/pull*, *come/go*.

4. Converses are included in the next category. These terms describe a relation between two entities from alternate viewpoints i.e. *above/below*, *own/belong to*. Palmer (1981:97) provides a similar taxonomy called relational opposites. Palmer suggests that, a different kind of 'opposite' is found with pairs of words, which exhibit the reversal of a relationship between items. For example, *buy/sell*; if John sells to Fred, Fred buys from John. Some antonyms referring to spatial position also belong in this category – *above/below*. Kinship terms also belong here. Many of them indicate not only the relationship, but also the sex of the person concerned – *father* is the male parent and daughter is a female child. But in this case reversibility is blocked. As Palmer says, if we say John is Sam's father, it does not entail that Sam is John's son – Sam could be his daughter. So we have pairs, which indicate the same relationship but a different sex: *father/mother, son/daughter.* So, if we say that John is Sam's brother, does not entail that Sam is John's brother (she might be his sister).

5. Saeed's (1997) last category includes taxonomic sisters. There are antonyms, which are at the same level in a taxonomy. Some of these taxonomies are closed, like days of the week (*Sunday*, *Monday*) and some are open, like the flavours of ice-cream. There is always a possible extension for open taxonomies. Someone can always come up with a new flavour and extend the taxonomy.

3.2.4 Meronymy

Meronymy describes a part-whole relationship between lexical items i.e. *cover* and *page* are meronyms of *book*. Saeed (1997:70) points out that meronymy can be identified by using sentence frames like X is part of Y, or Y has X – e.g. (1) 'A page is part of a book', or (2) 'A book has pages'. Meronymy differs from hyponymy in transitivity, in the sense that hyponomy is always transitive but meronymy may or may not be. Saeed provides the following examples: a transitive example is when we say *nail* is a meronym of *finger*, and *finger* of *hand*. *Nail* is a meronym of *hand* because we can say (3) 'A hand has nails'. A non-transitive example is when we say *pane* is a meronym of *window* (4) 'A window has a pane', and *window* of *room* (5) 'A room has a window'; but *pane* is not a meronym of *room*, since we cannot say (6) 'A room has a pane'.

3.2.5 Synonymy

Synonyms must have a significant degree of semantic overlap; for example, *truthful* and *honest* (Cruse, 1986:266). Synonymy refers to 'sameness of meaning' (Palmer, 1981). Palmer suggests that there are no real synonyms. There are no words that have exactly the same meaning. As Lyons (1968:448) mentions, it is undoubtedly true that there are few 'real' (or 'absolute') synonyms in language. Even though this statement is vague since Lyons does not specify the number of 'few real synonyms', it is reinforced by Saeed (1997) who also points out that true or exact synonyms are very rare.

Palmer (1981) provides five ways in which possible synonyms can be seen to differ: (i) First, some sets of synonyms belong to different dialects of the (same) language. For example, the term *fall* is used in the United States and in some counties in Britain, where others would use *autumn*. However, Palmer suggests that these words, which belong to different dialects of the same language, are like the translation-equivalents of

two different languages. Thus, "these groups of words are of no interest at all for semantics" (Palmer, 1981).

(ii) Secondly, there is a more challenging situation with the words that are used in different styles, differing in degrees of formality (e.g. *gentleman*, *man* and *chap*) indicating that *gentleman* is more formal than *chap*.

(iii) Thirdly, some words may be said to differ only in their emotive or evaluative meanings. For example, the terms *liberty* and *freedom*, *hide* and *conceal*, each implies approval or disapproval. These kinds of terms are used to influence attitudes.

(iv) Fourthly, some words occur only in collocation with other words. For example, *addled* occurs with *eggs* or *brains*. Palmer suggests that these are true synonyms, which differ only in that they occur in different environments.

(v) Fifthly, many words are very close in meaning, or their meaning overlap. In this sense, there is a loose sense of synonymy. Dictionaries exploit this kind of synonymy and they do not provide the precise connections between themselves.

3.2.5.1 Near-synonyms, partial-synonyms and conditions for absolute synonyms

According to Lyons (1995), we should not confuse near-synonyms (expressions that are more or less similar, but not identical, in meaning) with partial-synonyms (which are identical in meaning but fail to meet the conditions of what is referred to as absolute synonymy as defined below). Although Lyons insists that near-synonymy is not the same as partial synonymy, it should be noted that by his definition near-synonyms qualify as incomplete synonyms, and therefore as partial synonyms (Cruse, 1986:292).

As we mentioned earlier, there are few 'real' (or 'absolute') synonyms in language. Absolute synonyms should satisfy the following three conditions (Lyons, 1995:61-62):

(i) all their meanings are identical;

Lyons mentions that it is easy to show that *big* and *large* are not synonymous in all of their meanings and that they are partial-synonyms. If we consider the following sentences (1) 'They live in a *big/large* house', (2) 'I will tell my *big* sister', (3) 'I will tell my *large* sister' we notice that the sentence number (2) is lexically ambiguous, by virtue of the polysemy of *big* in a way that sentence number (3) is not. These sentences show that *big* has at least one meaning which it does not share with *large*. As Cruse

(1986:268), however, mentions it would be impracticable to prove that two items are absolute synonyms by this definition, because that would mean checking their relations in all conceivable contexts. This is theoretically impossible since the number of possible contexts is infinite.

(ii) they are synonymous in all contexts;

This condition refers to the set of contexts in which it can occur (its collocations). For example *large* is not interchangeable with *big* in the following sentence, (1) 'You are making a *big* mistake'. The sentence (2) 'You are making a *large* mistake' is meaningful and grammatically correct but it is collocationally unacceptable.

(iii) they are semantically equivalent (i.e. their meanings are identical) on all dimensions of meaning, descriptive and non-descriptive.

According to Lyons (1995:63), two expressions are descriptively synonymous if, and only if, propositions containing the one necessarily imply otherwise identical propositions containing the other, and vice versa. In this sense, *big* and *large* are descriptively synonymous in one of their meanings and over a certain range of contexts. So one cannot without contradiction simultaneously assert that someone lives in a *big* house and deny that they live in a *large* house. Lyons, also, mentions expressive meaning, which is one kind of non-descriptive meaning. This refers to the situation where two descriptively synonymous expressions differ in respect of the degree or nature of their expressive meaning. For example, the words *huge*, *enormous*, *gigantic* and *colossal* are more expressive of their speakers' feelings towards what they are describing than *very big* or *very large*. It is more difficult to compare the first set of words among themselves in terms of their degree of expressivity (Lyons, 1995:64).

As Cruse (1986:292) points out, Lyons's definition (iii) leaves open the question of how many dimensions of meaning there are, and how to determine whether two words are identical on any particular dimension.

3.2.5.2 Ways of testing synonymy

Synonymy can be tested in a number of ways. According to Palmer (1981), one way is to substitute one word for another. As we mentioned earlier, there are no two words having exactly the same meaning. However, some words are interchangeable in certain environments only. For example *deep* or *profound* may be used with *sympathy* but only *deep* with *water* (Palmer, 1981:91). But, as Palmer points out, this way does not give us similarity of meaning; it rather indicates the collocational possibilities of the words. Another way is to find the 'opposites' (the antonyms). If the two appear to have the same antonyms, this is a reason for treating them as synonyms (thus *superficial* is contrasted with both *deep* and *profound*).

3.2.5.3 Relation between synonymy and hyponymy

A very interesting relation exists between synonymy and hyponymy. Hurford and Heasley (1983) define hyponymy in such a way that synonymy counts as a special case of hyponymy. They provide the following rule. "If X is a hyponym of Y and if Y is also a hyponym of X, then X and Y are synonymous" (Hurford and Heasley, 1983:107). For example, the synonyms *mercury* and *quicksilver* are hyponyms of each other, and for this reason, synonymy can be seen as a special case of hyponymy, i.e. symmetrical hyponymy. This connection indicates the strong associations that can exist between these two different sense relations.

3.3 Conclusion

It emerges from the above discussion that there are different kinds of sense relations between words. We noticed that Hurford and Heasley (1983) identify polysemy when the same word (word form) has several very closely related senses (*mouth* of a river vs *mouth* of an animal). Homonyms are word forms pronounced or spelt in the same way but having unrelated senses far apart from each other and not obviously related to each other in any way (*mug* as drinking vessel vs *mug* as gullible person). Palmer (1981) mentions that when one form has several meanings, it is not always clear if this is a polysemic word or if this is a case of homonymy. Hyponymy refers to the notion of 'inclusion'. A hyponym includes the meaning of a more general word (*dog* and *cat* are

hyponyms of *animal*). Antonyms are words which are opposite in meaning. They are divided into binary pairs (*dead/alive*), gradable antonyms (*honest/dishonest*), reverses (*come/go*), converses (*above/below*) and taxonomic sisters (*Monday/Friday*). Meronymy describes a part-whole relationship between lexical items (*page* is a meronym of *book*). Synonymy is used to mean 'sameness of meaning'. Real synonyms are very rare. According to Lyons (1995), we should not confuse near-synonyms (expressions that are more or less similar, but not identical, in meaning) with partial-synonyms (which are identical in meaning but fail to meet the conditions of what is referred to as absolute synonymy). Lyons provides the following conditions for absolute synonyms: (a) all their meanings are identical; (b) they are synonymous in all contexts: and (c) they are semantically equivalent (i.e. their meanings are identical) on all dimensions of meaning, descriptive and non-descriptive.

The basic conclusion drawn from the above discussion is that we can state the meaning of words in terms of their associations with other words. The implication is that the study of sense relations between words can be applicable to vocabulary teaching. The sense relations described above can be useful not only for introducing new vocabulary to the students (i.e. presenting synonymous words together) but also for creating activities to enrich learners' better understanding of the meaning of the words they learn (i.e. being able to distinguish the difference in meaning between synonymous words).

It is worth mentioning that the three previous chapters provide a very useful ground for the present study described in Chapter Five. Mental lexicon and vocabulary development (Chapter One) presents information about the word form and meaning connection which will be tested in this study (see Chapter Five). Moreover, Chapter Two (learning and teaching L2 vocabulary) discusses a wide range of learning strategies and teaching methods some of which will be used in this empirical study. In addition, specific sense relations described in Chapter Three will also be used. The next chapter (Chapter Four) is considered to be an important part of the present research because it provides an account of two contrasting views regarding the use of semantic links (i.e. sense relations) or networks in classroom materials and activities for vocabulary learning in a L2.

CHAPTER FOUR

Two contrasting views of presenting and learning vocabulary

4.1 Introduction

In recent years, contradictory advice to teachers has emerged from studies into the use of semantic links or networks in classroom materials and activities for vocabulary learning in a L2. The literature presents two contrasting views: the first is that learning related words (see definition below) together at the same time makes learning more difficult, and the second is that presenting words of related meaning together makes learners see the distinctions between them and gain valuable knowledge of the defined area of meaning. These two ways of presentation may have value at different stages of learning.

By related vocabulary, I mean words that are related to each other in various ways. Words may be grouped in many different ways. This kind of word grouping is called clustering. According to Tinkham (1997:141), there are two manners of clustering: a) linguistically-based clustering and b) cognitively-based clustering.

a) Linguistically-based clustering

Linguistically-based clustering involves words in 'lexical sets'. A lexical set consists of vocabulary items which are grouped together because they share certain semantic and syntactic similarities (Crystal, 1997:221). The literature presents lexical sets as groups of words which are topic-related e.g. *apricot*, *plum*, and *peach*. This set of words is under the common superordinate covering topic (or concept) of fruits. These words share some aspects of meaning that are not present in the word *table*, for instance. All these words are consequently gathered together by virtue of their shared semantic and syntactic characteristics. Such vocabulary items are labelled as 'lexical sets' by some (Crystal, 1997), though others prefer the term 'semantic clusters' (Marzano and Marzano, 1988, Tinkham, 1993, 1997).

Schneider, Healy and Bourne (1998) present another label for semantic clusters. They talk about vocabulary grouped by 'conceptual category'. According to them, one

common method used to teach foreign vocabulary is to group together words from the same category. For example, students may be taught all the terms for the body parts in one lesson and all the terms for food types in another lesson. In this sense, all the terms for the body parts are grouped by conceptual category (Schneider, Healy and Bourne, 1998). Similarly, all the terms for the food types are vocabulary items which constitute a different conceptual category.

Another type of linguistically-based clustering is grouping words by sense relations. If we consider that semantic clusters consist of vocabulary items that are grouped together in sets of words sharing certain formal or semantic features, we may also consider that another way of grouping vocabulary items is by sense relations. This type of clustering could involve synonyms (*big/large)* or antonyms (*hot/cold)* because the words that belong in these groups share certain semantic properties (see Chapter Three).

b) Cognitively-based clustering

Cognitively-based clustering involves words that are subconsciously organized within certain 'frames' or 'concepts' which segmentize a speaker's background knowledge (Tinkham, 1997:141). There are two types of cognitively-based clustering: a) 'thematic clustering' and b) 'individually invented clustering':

a) 'Thematic clustering' is based upon psychological associations between clustered words and a shared thematic concept (Tinkham, 1993, 1997). A cluster of words drawn from such a frame might include *frog*, *pond*, *hop*, *swim*, *green* and *slippery*. These words of different parts of speech are all closely associated with a common thematic concept, in this case, *frog*.

b) The second type of cognitively-based clustering involves word-relations that are invented by individuals. I call this type of clustering 'individually invented' because a relationship between vocabulary items can be invented or created by any individual in their mind. For example, the pronunciation of the English word *terrace* /ˈterəs/ is almost the same as the pronunciation of the Greek word *τέρας* /ˈteræs/ which means *monster*. In this case, Greek EFL students could easily create a relation between the words *monster* and *terrace* and group these words together.

In this chapter, arguments for and against the presentation of related vocabulary in groups (or sets) will be discussed. The focus will be on four studies (Tinkham, 1993, 1997, Waring 1997, Schneider, Healy and Bourne, 1998 and, Finkbeiner and Nicol, 2003) which present similar evidence that it takes more time to learn related words together (at the same time) than it takes for learning unrelated words. The limitations of these studies and comments on the experimental designs of the experiments will also be considered.

4.2 Arguments for the presentation of related vocabulary in sets

Arguments for the presentation of related vocabulary in sets are mainly based on theory and not on experimental evidence. Nation (2000:6) provides five reasons for teaching related words in sets:

- it requires less learning to learn words in a set (Neuner, 1992);
- it is easier to retrieve related words from memory;
- it helps learners see how knowledge can be organized (Dunbar, 1992);
- it reflects the way such information is stored in the mental lexicon (see chapter one); and
- it makes the meaning of words clearer by showing how they relate to and are different from other words in the set.

Nation (2000) acknowledges the fact that it seems a good idea to present words of related meaning together so that learners can see the distinctions between them and gain a complete coverage of a defined area of meaning. He mentions that numerous writers (e.g. Channell, 1981, 1990; Neuner, 1992) suggest teaching related words in sets.

Semantic theory provides a systematic description of the vocabulary of a language. According to Channell (1981), we should teach L2 vocabulary in semantic sets (word groups sharing certain semantic characteristics) because the vocabulary of a language consists of interrelating networks of relations between words. These networks are called 'semantic fields'. The 'semantic field' theory suggests that the lexical content of a language is best treated not as a mere aggregation of independent words or an unstructured list of words but as a collection of interrelating networks of relations between words (Stubbs, 2001). Words belong to the same semantic field when they

share some aspects of meaning. At the same time, they hardly ever share all aspects. Synonymy (*large/big*), for example, is often confusing for the learners, since there are in practice very few words in any language which are interchangeable in all contexts (see Chapter Three). Similarly, 'componential analysis' offers a systematic and easy way of describing such similarities and differences. It consists of breaking down the meaning of a word or words into different pieces known as semantic components or features (Lyons, 1995). An analysis of walk, for example, might be: [+move] [+by feet] [+on land] [+placing down one foot after another] [+contact maintained with ground].

Channell (1981, 1990) argues that by analysing vocabulary into 'semantic fields', we are no longer dealing with random lists, but with a systematic structure. This is important because the mind uses semantic similarity in classifying words (Channell, 1981). She mentions that the mental lexicon takes account of semantic similarity in organizing words. For example, speech errors made by native speakers (slips of the tongue) show that wrong words come from the same semantic field as the intended word e.g. *I have my book and my jigsaw … I mean my crossword.* Thus, in terms of storing words into the mental lexicon, it is plausible to assume that a method of teaching that takes account of the semantic relatedness between words must be more effective than one that does not. It is therefore logical to explicitly teach some L2 vocabulary in semantic fields. Channell (1981) suggests the use of grids to tell the learner exactly what he needs to know about the relationships between words in the field, by making explicit their differences and similarities. For example:

	one's toes	one's fingers	one's fists	one's teeth
crack	+	+		
clench		+	+	+
grind				+

(Rudzka et al., 1985:118)

Channell (1981:121) also offers exercises like: *Choose from the words in brackets the one which best fits the given context: "As we reached the top of the hill a _____ view stretched out before us" (good-looking, handsome, beautiful).*

Channell (1981) points out that an important thing a learner needs to know about every new word he/she wants to make part of his/her active vocabulary is to know the way the new word relates to other words with similar meanings. Martin (1984:136) also mentions that L2 teachers, in presenting new vocabulary, can contrast new words with any synonyms appearing in the lesson or even, if the item is glossed with a synonym, determine whether the new item is a hyponym of the old one and supply the additional information that restricts the new one. Conscious awareness of the interrelationships among words provides learners with a tool that enables them to process input more effectively (Lewis, 1997:260), as well as a tool for organizing the mental lexicon (Singleton, 1999:273). In other words, as students come to notice the multiple factors in choosing words for contexts, "they will begin to ask the kinds of questions that will lead them to increased proficiency in expressing meanings fluently and accurately" (Martin, 1984:136).

4.2.1 Semantically related vocabulary in a classroom activity

In this section, I discuss a particular classroom activity which incorporates an explicit approach towards the presentation of semantically related vocabulary. I believe that the discussion of this activity presents a good argument for the presentation of related vocabulary in sets because it clearly shows how the theory of presenting semantically related items could be used in practice as a classroom activity.

Jullian (2000) describes an activity carried out with upper-intermediate and advanced learners to help them increase their word-meaning awareness and expand their active vocabulary. She introduced these activities in an optional course (Lexical Production course) for Spanish learners of English with a good command of English. She mentions that these learners manage to communicate in a satisfactory manner by using very limited lexical items. Their active English vocabulary is made up almost entirely of core words, the overuse of which makes their discourse uninteresting and sometimes childish. For example, they overuse words like 'nice' instead of trying 'friendly', 'delicious' or 'pleasing'. The study of semantic features in related words emerged as one of the components of the Lexical Production course. The activity described here helps students to understand the full semantic content of related words and to detect what

makes them similar and different from each other. According to Jullian (2000), this helps students acquire lexical semantic awareness. The objective is to analyse the meaning of close terms so as to detect what makes each word convey a unique meaning and what are the subtleties that make near-synonyms different from each other (Jullian, 2000). This contributes to acquiring a better mastery of them, given that to understand a word fully and use it appropriately "the learner needs to know [...] how it relates to others of similar meaning and which other words it can be used with" (Carter and McCarthy, 1988:49). In this activity, each word's paradigmatic and syntagmatic relations are considered.

Before I describe Jullian's task, I will first refer to the terminology used in the task. The way to penetrate the semantic content of words is by means of dictionary definitions. The semantic decomposition in Jullian's task is carried out following Hudson's (1995) 'steps in word-meaning' study. Students 'unpack' dictionary definitions of related terms into their constituent facts, which are usually built round a 'classifier' - which classifies a term in relation to a more general concept, and 'distinguishers' - which help to distinguish it from other hyponyms under the same classifier. He suggests that definitions reflect both similarities and differences between a word and other similar terms (Hudson, 1995:26). Hudson's 'classifier' corresponds to Jullian's *leading word*, which stands as the hypernym of the lexical set which comes after it, while his 'distinguishers' correspond to her semantic features, which make the elements of the set differ from one another. In other words, this activity aims at what Channell recommends: that vocabulary teaching should involve making associations between semantically related items (Channell, 1990:27).

The activity consists of 6 stages:

1) During the first stage, the students collect related words of a leading word (the classifier *hit*) to create a lexical set. The students suggest related terms, starting with those which are close in meaning to *hit*. Most of the students use the words *hit*, *strike* and *beat* confidently. Few of them would use the other words from the set with the same confidence. By using a dictionary, the students try to detect the semantic feature that makes each word unique, i.e. the way you hit, the place where you hit, the intention, etc.

2) In the second stage, the students classify all the words, according to common features, and create sub-categories depending on the features they want to emphasize. Some categories supplied by the students are the following: 1) once, 2) repeatedly, 3) with an instrument, 4) accidentally, etc.
3) During the third stage, the students draw a semantic network around the leading word to illustrate the categories they have created.
4) After dealing with paradigmatic relations, they move onto the syntagmatic relations. In the fourth stage, the students try to identify the context in which these words are likely to occur, and find appropriate collocations for them by making out suitable subjects, objects, adverbs, and complements for the words they are studying in isolation. They also study the grammatical use of each word; whether it takes an object or not, whether it can be used in the passive voice, etc.
5) In the fifth stage, the students are invited to write illustrative sentences in which the context will help to exemplify the meaning of the target word by providing as much information as possible about its content. For example, 'He *thumped* on the table in such fury that everything on it jumped and his fist started to swell'. Moreover, the students are invited to create sentences in which these words are used metaphorically.
6) Finally, the students have to choose a few of the most frequent and useful words to learn. The exercise concludes with the students writing a passage using the words they have learned from the set. In the final stage, the students are asked to do their own research and present it to the class.

Jullian (2000) mentions that the activity produced some very interesting results. Students realised that English is more synthetic and came to appreciate that their discourse gains in strength, accuracy, and beauty with the use of more precise words. This method helped students to become active agents of their own learning. The activities carried out contributed to the development of a wide range of lexical skills, a better understanding of their meaning, and a faster incorporation into their L2 mental lexicon (Jullian, 2000). According to Jullian, the students enjoyed the activity, while at the same time seeming to make considerable improvements in their competence.

4.2.2 A survey of L2 coursebooks

In addition to Jullian's classroom activity, further evidence for the presentation of related vocabulary in sets, and particularly in semantic clusters (topic-related vocabulary), comes from ESL textbooks. According to Tinkham (1993), ESL curriculum writers assume that the simultaneous presentation of semantically related lexical items somehow makes learning easier. Waring (1997) points out that it is a common practice in many current L2 coursebooks to introduce words in semantic groups. It appears that semantic clusters fit quite nicely into most current ESL textbooks. For example, learners are asked to learn 'foods' in *Headstart Beginner* Unit 5 (Beaven, 1995); 'jobs' in *Headway Elementary* unit 3 (Soars and Soars, 1993). There seems to be a belief among coursebook-writers that doing so will aid vocabulary building and lexical association in particular. According to both Tinkham (1993, 1997) and Waring (1997) this belief appears to be founded in methodology rather than on research.

Tinkham (1997) believes that two motivations appear to drive this clustering of semantically related vocabulary items:

1. First, semantic clusters service the methodologies driven by two current approaches to L2 development. For coursebook-writers of a structure-centred approach, semantic clusters fit naturally into 'slots' left open in the oral and written substitution activities. *New Horizons in English 1* (Walker, 1991:41) offers *carrots, nuts, grapes, pears, peaches, oranges*, and three more food labels as possible ways to complete the question, 'Do you like ____ ?'

   Following a more learner-centred approach to L2 development, Tinkham (1997) points out that other ESL programmes are guided by their writers' perceptions of the communicative needs of their students. As a consequence, these programmes are divided into various units responding to either situations in which students might find it necessary to communicate in their language (e.g. visiting a doctor), or dealing more closely with the language itself, the notions (expressions of time, location) and functions (e.g. requests) which students may wish to communicate.

2. The second reason which promotes the use of semantic clustering of new vocabulary items, according to Tinkham (1997), is the belief that such clustering facilitates the learning of new words. Gairns and Redman (1986:31) state that grouping words according to their semantic features can provide a useful framework for the learner to understand and see where meaning overlaps and to learn the limits of use of an item. Similarly, the concept 'spoon' is facilitated by learning how spoons are both similar to and different from knives and forks (Tinkham, 1997). Tinkham points out that although semantic clusters fit nicely into two current ESL methodologies and facilitate focus upon semantic similarities and differences among words being learnt, there is little direct empirical evidence that semantic clustering does in fact facilitate learning.

Let us now examine what is the typical way of presenting and explaining vocabulary in ESL coursebooks. Judging by a recent survey of ESL textbooks, it appears that many if not most ESL /EFL students are often exposed to their new language vocabulary preorganized for them in semantic clusters (topic-related vocabulary). The coursebooks that have been selected for this survey are for beginners, intermediate, advanced and proficiency students. The main interest is to examine the vocabulary section or task in each unit of the coursebooks. Follow-up exercises will also be taken into account. The main focus will be on examining the vocabulary presentation in terms of relatedness of meaning or not. I will not examine other aspects of vocabulary presentation such as collocations and fixed expressions, connectors, adverbial phrases, phrasal verbs or idioms because my main priority is to focus on single vocabulary items and their meanings. The following table (Table 4.1) includes all the L2 coursebooks used in this survey and provides an idea of how vocabulary is presented in these coursebooks.

Table 4.1: Vocabulary presentation in L2 coursebooks

Coursebook Title	*Level of Students*	*Is Vocabulary Related by Topic?*	*Is Vocabulary Not-related?*	*Exploitation of Synonyms, Antonyms and Homonyms*
Headstart	*Beginner*	*Yes*	*No*	*No*
Headway	*Elementary*	*Yes*	*No*	*No*
Cutting Edge	*Starter*	*Yes*	*No*	*No*
English File 1	*Beginner*	*Yes*	*No*	*No*
Cutting Edge	*Intermediate*	*Yes*	*No*	*No*
Inside Out	*Intermediate*	*Yes*	*No*	*No*
Matters	*Intermediate*	*Yes*	*No*	*Yes*
Cutting Edge	*Advanced*	*Yes*	*No*	*No*
Gold Proficiency	*Advanced*	*Yes*	*No*	*No*
Knockout FC-Workbook	*Advanced*	*Yes*	*No*	*No*
Countdown to First Certificate	*Advanced*	*Yes*	*No*	*No*
English Vocabulary in Use	*Advanced*	*Yes*	*No*	*No*
The Cambridge English Course 3	*Advanced*	*Yes*	*Yes*	*No*
Wordwise	*Advanced*	*No*	*Yes*	*Yes*
Headway	*Advanced*	*Yes*	*No*	*Yes*
Gold Proficiency	*Advanced*	*Yes*	*No*	*Yes*
Plus FC	*Advanced*	*Yes*	*No*	*Yes*

There are many examples found in the coursebooks that present new words in lexical sets based on topic-related vocabulary. *Cutting Edge Starter* (Cunningham, Redston and Moor, 2002) provides many examples of topic-related vocabulary. For example, in Module 4 (p. 30), the vocabulary is about places in a town e.g. a *bank*, a *post office*, a *cinema*, etc. Moreover, on p. 57 the days of the week are presented together. In addition, the vocabulary section on p. 100 is about colours and sizes.

In *English File 1* (Oxenden and Seligson, 1996), words again are presented in topic-related groups: the numbers are presented together on p. 12 and phrases for giving directions are also grouped together on p. 84. In *Cutting Edge Intermediate* (Cunningham and Moor, 1999), the vocabulary section in Module 1 is about people around you (p. 8). The exercise asks students to check the meaning and the pronunciation of the words in the box and then write the words in the correct column below. The vocabulary is topic-related. It is about people around you e.g. *best friend*, *cousin*, *classmate*, *stranger*, etc.

Inside Out (Kay & Jones with Hird & Kerr, 2001) also uses topic-related vocabulary. The vocabulary items in Unit 5 are topic-related (about football). *Intermediate Matters* (Bell and Gower, 1995) also presents related words together in lexical sets. In the beginning of Unit 7 (p. 49) with the title "*Choosing a partner*" there are two boxes which contain words and phrases about personality (*confident, generous, patient, ambitious, a sense of humour,* etc.) and appearance (*pretty, elderly, well-built, balding,* etc.). Later on, in the same Unit, in the Vocabulary Section, Exercise 1 asks students to complete the gaps in the dialogues with words from the boxes A and B which again contain words which relate to the general theme of Symptoms (*diarrhoea*, *pain*, *sick*, etc.) and Illness (*heart attack, flu, infection*, etc.).

In *Cutting Edge-Advanced* (Cunningham, Moor and Comyns Carr, 2003), Module 7 also presents words and phrases in groups associated with describing characteristics (*negative attitude, chatty, grumpy and irritable, unpredictable,* etc.). Other examples come from *Gold Proficiency* (Newbrook and Wilson, 2000) where Exercise 3.1 (p. 63) presents together the words *commit, plead, pass, convict/acquit, admit, serve, give, cross-examine, appeal, sentence, remand, charge,* all related to crime. *Knockout* FC – *Workbook* (Martin and May, 1999) also presents new vocabulary in lexical sets; for example, on p. 30 (Unit 5), the exercise presents the following words together: *palace, shed, stable, castle, bungalow, gallery, mansion, hut, attic,* and *chalet*. The topic of the Unit is *Home and Away* and these words reflect the theme of different houses and buildings being used as homes.

Moreover, in *Countdown to First Certificate* (Duckworth and Gude, 1999) in *WORDPOWER* section (Unit 9), the exercise asks students to match the words with their meanings. The words presented here are all connected with talking (*gossip, mumble, whisper, yell, argue, discuss*).

In some cases, teaching vocabulary in lexical sets is explicitly stated. *English Vocabulary in Use* (McCarthy and O'Dell, 1994) was written to help students improve their English vocabulary and help them learn not only the meanings of words but also how they are used. It is explicitly stated in the book that students can help themselves to learn vocabulary by learning associated (related) words together (p. 4). One of the ways

stated in the coursebook is to learn words together that are associated in meaning. It is a popular and useful way of organising the vocabulary study. The coursebook-writers suggest using words in a network e.g. *cat, paw, kitten.* Another example comes from *Countdown to First Certificate* (Duckworth and Gude, 1999). The Teacher's Book explicitly recommends teachers to encourage their students to record new items of vocabulary in lexical sets for easier reference (p. 8).

The examples presented above clearly demonstrate that, in most cases, new vocabulary is presented to students in lexical sets. However, it is equally important to mention that there are some examples where vocabulary is presented in unrelated sets. In most cases, this fact has to do with vocabulary revision exercises. For example, in *Wordwise: a Proficiency vocabulary exercise book* (Philippakis, 1995:14), students have to choose the appropriate word or phrase from the box. The words are not related to each other (e.g. *exclusive*, *racial*, *beat*, *pasta*, *utter,* etc).

There are some books with exercises that exploit synonyms, antonyms or even homonyms. On p. 66 (*Intermediate Matters*) synonyms and antonyms are explicitly presented. The exercises in this section explain to the students what the terms *synonym* and *antonym* mean, and ask students to find antonyms in the box for each of the words in italics in order to complete the sentences e.g. I only drink *dry* wine. I hate_____ wine. The students are also asked to find synonyms for the adjectives in the following phrases e.g. an *attractive* man, a *nice* time, etc. An exercise to find synonyms and the opposites of words is found in *Wordwise: a Proficiency vocabulary exercise book* (Philippakis, 1995: 111). The exercise asks students to organize 45 given words in 15 groups, with each group containing: one word, one synonym of that word and one opposite of that word (e.g. *to desert, to abandon, to join*).

Headway-Advanced (Soars and Soars, 1989:16) also exploits synonyms and their associations. One exercise asks students to find synonyms or near-synonyms for words like *enemy*, and write sentences to explore their associations. In addition, the *Teacher's Book* urges teachers to show students that words are not isolated and unrelated but can be part of a pattern. Examples of these patterns are synonyms which, in this coursebook, are given particular attention.

Gold Proficiency (Newbrook and Wilson, 2000) also states that students at this level are expected to know the differences between words with similar meaning. Unit 9 (p. 129) draws students' attention to the difference in meaning between *heal*, *treat* and *cure* by providing their dictionary definitions. There is a follow-up exercise to ensure students are able to distinguish the difference in meaning between a *vision*, a *dream* and an *ideal* or between *romanticism*, *escapism* and *idealism*.

Moreover, in *Plus* (FC Level, by Moutsou and Parker, 1998) Unit 2, on p. 33, the writing practice exercise asks students to read the pair of sentences and replace the words in bold type with two of the words given. In this exercise the students are expected to know the differences in meaning between words such as *observed*, *stared*, *glanced*, *watched* and *looked*. These are words with similar meaning. This kind of exercise emphasizes the need for students to be able to distinguish differences between words with similar meaning.

4.3 Evidence against the presentation of related vocabulary in sets

While Nation (2000) finds it a good idea to teach related words in sets, he also refers to a growing body of research which shows that it takes more time to learn words that relate to each other in groups than it takes to learn words that are unrelated to each other. Tinkham (1993) found that presenting L2 students with their new vocabulary grouped together in lexical sets of syntactically and semantically similar new words might actually impede rather than facilitate the learning of the words. By learning, here, I mean the ability to recall/retrieve the meaning of previously met vocabulary items (see section 2.2.3). In his two experiments, Tinkham (1993) compared the learning rates of subjects (American University students) learning semantically related and semantically unrelated new L2 words. His findings suggest that students have more difficulty (it takes them more time) learning new words presented to them in semantic clusters/lexical sets (e.g. *jacket*, *shirt*, *sweater*) than they do learning semantically unrelated words.

According to Tinkham (1993), a very large body of evidence suggests that the learning of a particular word may interfere with the learning of other words which are semantically similar. 'Interference Theory' states that when words are being learned at

the same time, but are too 'similar' or share too many common elements, then these words will interfere with each other, thus impairing retention of them. Extensive research into interference theory (see Baddeley, 1990) suggests that as similarity increases between targeted information and other information learned either before or after the targeted information, the difficulty of learning and remembering the targeted information also increases (Tinkham, 1993). In other words, this theory suggests that if new words are to be presented to learners, they should not be presented in groups that share a common head-word or superordinate concept. For example, 'clothes' words such as *jacket*, *shirt* and *sweater* should not be presented in groups because the learner (depending on his/her level and education experience) will confuse them. Similarly, recent researchers have posited a 'distinctiveness hypothesis' (see Hunt and Mitchell, 1982), which relates ease of learning to the distinctiveness (non-similarity) of the information to be learned. The data collected by these researchers suggest that the presentation of new vocabulary items to L2 learners in clusters of semantically and syntactically similar nondistinct words impedes rather than facilitates learning.

The possibility, therefore, arises that most L2 students are struggling to learn new words which have been selected and presented to them in a manner that impedes learning, and that a different manner of selection and presentation might actually make learning easier. In an attempt to explore this possibility, Tinkham's studies (the first in 1993 and second in 1997) examine the effects of semantic and thematic clustering upon the learning of new L2 words. In the first study (1993), Tinkham examines only semantic clustering, while in the second study he examines both semantic and thematic clustering. The semantic and thematic clustering is intended to distinguish between two different manners of organizing lexical items. As mentioned earlier, semantic clustering is based upon semantic and syntactic similarities among clustered words [*apricot*, *peach*, *plum*, *nectarine*, *pear*, *apple*] and thematic clustering [*frog*, *pond*, *swim*, *hop*, *green*, *slippery*] is based upon psychological associations between clustered words and a shared thematic concept (Tinkham, 1993, 1997).

4.3.1 Tinkham's first study (1993)

In his first study (1993), Tinkham conducted two experiments. In both experiments, subjects listened to lists of English words paired with imaginary words, which comprised a set of words. The English half of each word pair was presented in a mixed order and the subject had to remember and say the imaginary (L2) half of the pair within a set time. Their task was to try to learn the word pairs in as few trials as possible. The criterion for learning was met when a set of words had been learned. Tinkham's study intended to test the hypothesis that L2 learners learn new words in their new language more easily (may take less time to retrieve the words) if those words are grouped together in sets of unrelated words, than they do if the new words are presented to them in semantic clusters (sets of semantically and syntactically similar words).

The aim of the first experiment was to determine if subjects learned three word-pairs, pairing semantically related English words with artificial words, more slowly than they learned three word-pairs pairing unrelated English words with similar artificial words (see below).

The word-pairs for experiment 1 were:

shirt – moshee	
jacket – umau	related
sweater – blaikel	
rain – achen	
car – nalo	unrelated
frog - kawvas	

The subjects learned the six word-pairs through multiple exposures to the pairs mixed together, in varying orders, in six-pair sets. Tinkham (1993) had a group of 20 subjects ranged in age from 16 years to mid-forties who were, for the most part, students at a large university in the United States. A trials-to-criterion test was prepared to determine how easily subjects could learn the word-pairs in a set of pairs, each pair consisting of an English stimulus word paired with an artificial 'new' word. Within the set, three of the English words were semantically related one to another, while the other three were

unrelated to any of the other words. The English words were all nouns that were well within the vocabularies of the subjects. The artificial words were created by the researcher after following certain phonological guidelines (see below) in order to decrease the possibility that a particular set of artificial words might somehow be more learnable than the others:

- All six artificial words have two syllables.
- Three of the six words receive stress on the first syllable while the others receive stress on the second.
- Three words end in a vowel.
- Two words begin with a vowel.
- One word contains a consonant cluster.
- One word contains a vowel diphthong.

(Tinkham, 1993:374)

The use of nonsense (artificial) words should be seen as a positive feature that increased the internal validity of the experiments by allowing the experimenter to control the meaningfulness of the forms.

The researcher altered the ordering of the pairs through the trials so that they all fell in the various positions possible within the set. The subjects were not exposed to the same pair twice consecutively because the ordering ensured that the last word-pair on one trial did not become the first pair on the next trial.

Two forms of the test were created. Form A presented the pairing of English word/artificial word, while form B presented a reversed pairing. That is, the artificial words originally paired with related words in form A were paired with unrelated words, and vice versa, in form B. The test was administered individually and orally. The trials of the test were presented to the subjects via a recording played by a tape and the subjects responded by saying the appropriate artificial words. The subjects first heard an initial modelling of each of the artificial words coupled with its corresponding English word, '*Moshee* means shirt'. These pairings were presented in no particular order. Following the initial modelling, the subjects participated in a succession of trials, the trials presenting, one by one, the six English words, each followed by a pause during which the subjects attempted to remember and say the corresponding artificial word. For each item within a trial subjects heard the English word followed by a 3-sec pause, the sound of a bell, and then the appropriate artificial word. The subjects were required to

say the appropriate artificial word during the pause. There was a 2-sec pause between items and a 5-sec pause between trials. There was a gap of two seconds before the next pair until all words in a set of six word-pairs had been met. This set of six word-pairings constituted one trial. For example, when the subjects heard 'shirt' (the L1 word), there was a gap of three seconds before which the subject was required to remember and say the L2 artificial word (in this case 'moshee'), then they heard the sound of the bell and then the corresponding L2 artificial word 'moshee' was given. The criterion for a condition was met when a subject correctly said (before the bell), during a single trial, each of the three artificial words included within that condition. When the criterion had been achieved for one condition, that trial was silently marked by the researcher and the test continued with all six word-pairs until the criterion was met for the words comprising the other condition. Exact pronunciation was not demanded as some of the subjects could not easily distinguish some sounds.

Tinkham (1993) found that the subjects learned three word-pairs pairing semantically related English words with artificial words more slowly than they learned three word-pairs pairing unrelated English words with artificial words, even though the related and unrelated pairs were mixed together in a series of six-pair trials.

Tinkham's (1993) second experiment was intended to determine if a group of subjects learned a set of six artificial words paired with semantically related English words with more difficulty than they did a separate six-word set of artificial words paired with unrelated English words. The same subjects were used. The tests providing data for experiment 2 were administered in the same manner as the test in experiment 1, the only difference being that experiment 2 consisted of two six-pair tests.

The word-pairs for experiment 2 (separate, related and unrelated) were:

Related	Unrelated
pear – okess	mountain – awnai
apple – nuga	shoe – tosel
apricot – beloot	flower – manzeek
plum – kaisher	mouse – kunop
peach – eckly	sky – efoo
nectarine – depai	television - chengee

He found that the subjects learned a six-word set of artificial words paired with semantically related English words more slowly than they learned a set of artificial words paired with unrelated English words.

The results of this study are significant considering that, while the two experiments arranged the word-pairs learned by the subjects differently (one mixed group of pairs in the first experiment; two homogenous groups in the second experiment), the results of the two experiments were remarkably similar.

During the post-test interviews, where the subjects were asked to tell the experimenter the learning strategies they employed in order to learn the new words they were presented, each subject was asked which of the three sets of words were the most difficult. Almost all the subjects responded that the related words were the most difficult because they could not think of any mnemonic word associations for the words. Many subjects found the related words confusing because the items were too similar.

4.3.2 Tinkham's second study (1997)

Tinkham's (1993) results suggest that new words are learned more easily if they are not grouped for presentation in semantic clusters. These findings suggest that perhaps we should not give wordlists to our learners which have words that come from the same semantic set, but should be asking them to learn words semantically unrelated to each other.

The same results appear in Tinkham's (1997) later study. The intention of this research was to investigate the effects of both semantic and thematic clustering upon the ease with which new L2 vocabulary items are learned. All the subjects (university students) were native speakers of English. Experiment 1 consisted of four separate studies, two conducted in the oral modality (responding to a word they heard by saying their response) and two conducted in the written modality (responding to a word they saw by typing their responses). Half of the students took tests orally and the other half took tests in the written modality. Of the two studies conducted in a particular modality, one study involved the task of recognizing new artificial words and the other involved the task of

recalling artificial words given their English meanings. The purpose of each study was to compare, by the independent variable 'condition', the learnability of three-pair sets of associate pairs pairing English and new artificially created words. The independent variable was nominally divided according to the following classification. Examples for each category are provided in the next paragraph:

- *Condition 1*: linguistically related 'semantic clusters': words of the same form-class which directly descend as co-ordinates under a common superordinate concept.
- *Condition 2*: linguistically unrelated sets: words of the same form-class which do not directly descend from a common superordinate concept.
- *Condition 3:* cognitively associated 'thematic clusters': words of different form-classes that, in accordance with the intuition of the author, were likely to be associated with a shared thematic concept.
- *Condition 4*: cognitively unassociated sets: words of different form-classes that were judged not to be associated with a shared thematic concept.

(Tinkham, 1997:143)

The materials for each of the oral studies consisted of two trials-to-criterion tests each involving a six-pair set of associate pairs pairing an English word with an artificial word. Each six-pair set included three English words from one condition and three from the opposing condition. The six word-pairs were mixed together in a variety of orders across trials. Both the recognition and recall studies consisted of one six-pair set which included three semantically related and three unrelated English words and another six-pair set which included three thematically associated and three unassociated English words. The English words, arranged by condition (see above), were:

Condition 1, semantic clusters		Condition 2, unrelated sets	
dish	shirt	acid	island
bowl	jacket	smoke	potato
plate	sweater	roof	beard

Condition 3, thematic clusters		Condition 4, unassociated sets	
beach	library	fork	triangle
sunny	whisper	count	improve
swim	quit	brave	sweet

The artificial words were created by the researcher following the same phonological guidelines as in his previous study (1993). The researcher altered the ordering of the pairs throughout the trials. Each subject was required to hear and recognize the artificial words and say the corresponding English word on two of the tests and hear the English words and recall (say) the corresponding artificial word on the other two tests. Thus, the 'task' was a within-subject independent variable. Half the subjects learnt one particular pairing of English and artificial words (Form A) while the other half of the subjects learnt a different pairing (Form B). For both forms, the English and artificial words were the same, the only difference being that the English and artificial words were paired differently for the two forms.

The English/artificial word-pairs employed in the two studies administered in the written modality were the same as those administered in the oral modality with some additional constraints:

- All words must have five letters.
- No words may contain double letters.

Subjects were tested individually in two different sessions separated by approximately two weeks. The first session involved the production of English words (recognition of new words) while the second session involved the production of the artificial words (recall of new words). Both sessions included the two tests from Experiment 1 as described above followed immediately by the four tests from Experiment 2 as described below. Two trials-to-criterion tests were individually administered to 24 subjects hearing artificial words and responding by saying the English words which represented the meanings of those new words. One of the tests mixed three semantically related English words with three unrelated English words while the other test mixed thematically associated words with unassociated words. The tests were administered via a recording played on a tape-recorder. For each item subjects first heard the artificial word followed by a two-second pause, the sound of a bell and finally the appropriate English word. Subjects were required to say or at least begin saying the appropriate English word during the two-second pause. Items within a trial were separated by a four-second pause and trials within a test were separated by a six-second pause. The two recall tests given in a second session two weeks after the session involving the

recognition tests, resembled the recognition tests, except that subjects learned new sets of English artificial word-pairs. The subjects were given the English words as stimuli and responded by saying the artificial words during a three-second pause. The tests administered in the written modality paralleled those in the oral modality but were administered visually via PC computer. The tests employed the same materials as well as the same overall procedures.

The subjects, procedures and design for Experiment 2 were the same as those for Experiment 1. There was only one difference between the two experiments. In Experiment 1 the six-pair sets included a mixture of three English words representing one condition and three representing another. However, in Experiment 2 the six-pair sets were constructed homogeneously. This means that each set was employing English words representing only one particular condition.

According to Tinkham (1997), Experiment 1 clearly indicates that new L2 vocabulary items arranged in semantic clusters are learnt with more difficulty than new vocabulary items learnt in unrelated sets. This means that the subjects took more time to retrieve the new L2 artificial words arranged in semantic clusters. He also found that the same results appear in the data from Experiment 2. The data also indicated that new L2 vocabulary items arranged in thematic clusters are more easily learnt than new L2 vocabulary items arranged in unassociated sets.

4.3.3 Waring's study (1997)

Waring's (1997) study is a close replication of Tinkham's (1993) study. He also wanted to see if the same effects occurred with Japanese subjects rather than conducting the experiment in English as Tinkham had done. If the same effects were found for subjects with a different L1, then the results may demonstrate some generalizability to other languages. His subjects were eighteen native-speaking Japanese and two non-natives with advanced proficiency in Japanese. Most of the subjects were studying at educational institutions.

Waring's (1997) and Tinkham's (1993) studies conclude that presenting students with wordlists of new words in semantic clusters, rather than in unrelated word groups, can interfere with learning. The format of Tinkham's study was followed as closely as possible using the L1 words he selected for his work. Except that they were translated into Japanese. But Experiment 2 presented a problem because the fruits chosen for the original study are not typical examples of fruit in Japan. For this reason, several native speakers of Japanese were interviewed to determine the most typical examples of fruits one would find in Japan and Tinkham's list of fruit was revised in light of this.

In Waring's study, new words were made up following the guidelines in the original study because some of the artificial (L2) words in the original study were very close in spelling and pronunciation to existing Japanese words. For example, 'kaisher' sounds like the Japanese word 'kaisha' meaning 'company'. Experiment 1 consisted of two sets of three word-pairs (as in Tinkham).

The word-pairs for experiment 1 were:

L1	L2	
shatsu (shirt)	kilme	
jyaketo (jacket)	ifpa	related
seta (sweater)	blaikel	
ame (rain)	uchen	
kuruma (car)	nalo	unrelated
kaeru (frog)	kawvas	

Experiment 2 consisted of two sets of six word-pairs to be learned. One set of words shared a common superordinate concept (fruit) and was labelled the 'related words' and the other six words did not share a common superordinate concept and were labelled the 'unrelated words'.

The word-pairs for experiment 2 (separate, related and unrelated) were:

Related		Unrelated	
L1	L2	L1	L2
meron (melon)	ijos	yame (mountain)	ejaut
ringo (apple)	denga	kutsu (shoe)	tostrel
ichigo (strawberry)	esmek	hana (flower)	padeen
budoo (grape)	pairnya	nezumi (mouse)	kunop
momo (peach)	uldon	sora (sky)	efoo
mikan (orange)	nakew	terebi(television)	chengee

Waring (1997) reports that the results from both experiments show that presenting new words that share a common superordinate in a set of words to learn does interfere with learning.

During the post-test interviews (as in Tinkham's research), the subjects were asked to tell the experimenter the learning strategies they employed in order to learn the new words they were presented. Most of the subjects reported using a mnemonic device to try to remember the words (e.g. visual image). Waring reports that some learners found the unrelated set of words easier to learn because many of the objects were in the room or outside the window of the room where the data were being collected. However, no one commented on the ease of learning of the 'clothes' words despite the ready availability of sweaters and shirts (Waring, 1997).

4.3.4 Major findings of the three studies

The major finding of the three studies mentioned above is that semantic clustering of new L2 vocabulary items appeared to serve as a detriment to learning. In two separate but parallel experiments yielding similar results, sets of artificial words paired with semantically related English words (e.g. *apple*, *pear*, *nectarine*, *peach*, *apricot*, *plum*) were learnt with more difficulty than artificial words paired with sets of unrelated English words (e.g. *paint*, *funeral*, *recipe*, *market*, *uncle*, *ice*). Tinkham (1997, Experiment 2) found that, of 96 possible individual comparisons involving semantically related lexical sets and unrelated words, learning was faster for unrelated words for 80 comparisons, there was no difference for 13, and learning was faster with related words for 3. Once again, the effect is strong.

For thematically related words and unrelated words, the difference is not so marked. The thematically related words took 15% fewer repetitions to learn than the unrelated words. We have this result partly because the thematically related words were made up of different parts of speech (Nation, 2000). As mentioned, the semantically related lexical sets were all nouns whereas the thematically related set contained nouns, verbs and adjectives. Nouns are generally easier to learn than verbs, adjectives or adverbs (Laufer, 1997).

Nation (2000:7) suggests that if learning related words together has a small interference effect on learning, then it is not worth making any changes to the way words are grouped for learning. If the interference effect is large, however, the teachers and learners need to try to reduce the possibility of interference. One way is to learn related words separately. Tinkham and Waring (1997) found that the strongest interference effect occurs when all the words in a group to be learned are related to each other. They found that it took from 47% to 97% more repetitions to learn the groups of related items, as compared to the number of repetitions it took to learn the group of unrelated items (Nation, 2000). These are large differences. As both researchers point out, the results of these studies would probably come as a surprise to many current writers of ESL texts who, for a variety of reasons rely heavily upon the employment of semantic clusters in their presentation of new vocabulary.

4.3.4.1 Criticisms of Tinkham's (1993) and Waring's (1997) research

Waring (1997) discusses some limitations on the generalizability of the results found in both studies (Tinkham's first study in 1993 and Waring's study in 1997). I will present these limitations as they appear in Waring's study.

- Waring points out that the researchers have found that **the occurrence of interference depends on the type of stimulus material**. When meaningful passages are used rather than lists of words or nonsense syllables, no interference effects are found (Haberlandt, 1994:211). In Tinkham's (1993) and Waring's (1997) studies the words were learned in lists and this effect may not hold for words learned from a meaningful passage.

- Second, **very few words were tested**. Additionally, the words were learned orally and the effect may not hold for learning from written information. This can be tested by asking subjects to write the words rather than only say them.

- Third, the testing was on the productive use of the words. Another experiment could test if the effect also occurred receptively. Finkbeiner and Nicol (2003) (see section 4.3.6) provide an answer to this question. The results of their study revealed that participants translated L2 labels learned in semantic sets significantly more slowly than they did L2 labels learned in random order. This was the case in both (L1-L2 and, L2-L1) translation directions.

- Fourth, there are limits on the trials-to-criterion method whereby a condition was met when all the words in a semantic set had been produced correctly in one trial. Waring (1997) mentions that as the learner was trying to learn the other set, some of the first set, which had already been checked by the researcher as learned, were forgotten temporarily. Sometimes the subjects were surprised when the session ended with all the words being checked as learned as they had not felt they knew all the words properly and may have needed one or more trials to be sure.

- Fifth, **the artificial words that were used were strictly controlled**. This means that one cannot generalize these findings to natural languages without qualifications being made. This is especially true because the words were chosen to counterbalance effects for word shape. However, this created a problem in that the artificial words are less homogenous, in terms of graphotactics and length distribution, than the Japanese ones. In future studies where artificial words are being used, in order to avoid these problems, researchers should strive to create words with a similar consonant-vowel structure to the L1 against which they are being tested. It might be possible to use a natural language of which the subjects had no knowledge. However, according to Waring (1997), this creates problems and a language would need to be found where the learning for each of the words would be similar, in terms of words with different stress patterns, the number of syllables, the prototypicality of these words in the L1 and so on. In

other words, the researcher should balance the words against each other for learnability effects.

- Sixth, it seems that **there is no clear definition of what semantic relatedness might mean**. In this experiment words were chosen to show unrelatedness. It is clear that some words fit neatly into 'closed' sets, such as days of the week, but there are more 'open' sets that can have rather looser borders. For example, does one classify an electric mixer, a knife sharpener, an egg slicer or a cutting board as 'kitchen utensils'? There is a need for clear definition of terms before research is commenced (Waring, 1997).

- Seventh, according to Waring (1997), it is not clear that this trials-to-criterion measure is so straightforward. The number of learning trials in the second experiment is fewer than for the first experiment despite the subjects having to learn the same number of words in both. It may be that task-learning effects were affecting the data. That is, the subjects got better at doing this kind of task as the experiments progressed (practice effect).

- Lastly, **it is not clear whether these same effects will hold for learners who already have part of the semantic set being tested**. For example, if learners already knew 10 words from the 'clothes' semantic set and were being asked to learn some more words they would be adding to, rather than setting up, a new semantic set in the L2 (assuming the 'clothes' semantic network in the L1 was unavailable in the L2). That is, the learner does not already have a target language network set up to add the new words to. Waring (1997) points out that the effects found in this study may therefore be restricted to beginning learners rather than intermediate ones as the beginning learner has to set up semantic and vocabulary knowledge networks in the L2 into which the words must be put. An intermediate learner would probably already know many words from the semantic groups and when presented with new words may only need to add new words to an existing store, rather than create a new one from scratch. Nation (2000:6) also mentions that learning related words in sets is not a good idea for initial learning. As learners' knowledge becomes more established, seeing

related words in sets can have a more positive effect. According to Nation (2000), this question awaits a future study.

4.3.4.2 Comments on the design of the experiment

Waring (1997) also make some comments on the design of the experiment. He points out that the experimental design was stressful for the subjects as they were constantly under time pressure and their thinking was interrupted. Larger word sets would have taken too long to administer in one experiment. There were also problems with consistency of marking. The researcher had trouble assessing whether a word had been correctly produced. There was very little time for the researcher to assess a word and score it correctly, as assessment was done concurrently. It was also difficult to determine whether a word had been correctly supplied before or after the bell as sometimes both occurred simultaneously.

Moreover, Waring (1997) refers to the nomothetic research tradition within his and Tinkham's (1993, 1997) work. Ochsner (1979) made the distinction between *nomothetic* and *hermeneutic* traditions of inquiry. The dominant experimental or nomothetic pre-paradigm that exists in SLA attempts to explain the classroom, learners and learning from the results of experiment. The intention of experimental or quantitative research is to look for a single reality or truth. The qualitative or hermeneutic research tradition seeks to discover about the classroom in naturalistic, interpretive or qualitative terms allowing for multiple realities. Meara (1996:38-39) exemplifies the nomothetic view by saying that we need a "challenging combination of real-world constraints and rich theory". In my opinion, we need a balance between the dominant nomothetic tradition and the hermeneutic tradition. At the moment, the relations between the two are poor, and the coursebook-writer does not know which side to listen to. Tinkham's study clearly fits squarely into the former paradigm. According to Waring (1997), as the variables were tightly controlled in this experiment, it renders them somewhat ungeneralizable for our classrooms. In other words, the results these experiments generate might not fully apply to the natural environment of a classroom. Thus, the experiments lack external validity.

4.3.5 Schneider, Healy and Bourne's study (1998)

Schneider, Healy and Bourne's study (1998) initially appeared to suggest that learning related words together (e.g. parts of the body) was easier than learning unrelated words. In practice, this means that the participants (see below) took less time to retrieve the L2 words. However, when a test of long-term retention was administered the results did not appear to be the same. Nation (2000:7) points out that this research used a different way of choosing unrelated words from Tinkham's (1993) study, and this may have obscured some of the differences between the related and the unrelated groups. What makes the findings of this study interesting enough is the fact that, contrary to the above researchers, Schneider, Healy and Bourne used natural-L2 words and not artificial words. They conducted two experiments in order to investigate the effects of interference on learning and retention of foreign vocabulary.

Experiment 1

In the first experiment, participants (twenty-four non-French-speaking college students) learned the association between French words and their English equivalents, with the words either grouped (or blocked) by conceptual category (lexical set based on topic-related vocabulary) or presented in a mixed order (unrelated vocabulary). In the acquisition phase, 25 French-English word pairs were presented to participants on a computer screen in groups of five, at the rate of two seconds per pair. After a given block was shown for study, participants saw the five French words one at a time and were asked to type the English equivalent. For the participants in the grouped-order acquisition phase, the five words in each group were related. For example, one group of words consisted of body parts: *dos*, *back*; *bouche*, *mouth*; *figure*, *face*; *doigt*, *finger*; *yeux*, *eyes*. For participants in the mixed-order acquisition phase, the words in each group were unrelated and consisted of words from each five categories. For example, one group was *dos*, *back*; *avion*, *airplane*; *assiette*, *plate*; *jambon*, *ham*; *chemise*, *shirt*. The participants were next presented with another block of five pairs, and so on until all word pairs had been presented and tested. The following table (Table 4.2) depicts the distinction between the grouped-order condition and the mixed-order condition.

Table 4.2: Grouped-order condition and the mixed-order condition (Schneider, Healy and Bourne's study (1998)

	Grouped-order condition	**Mixed-order condition**
Definition	A group of words based on topic-related vocabulary	A group of words based on unrelated vocabulary
Examples	*dos,back*; *bouche,mouth*; *figure,face*; *doigt,finger*; *yeux, eyes*	*dos,back*; *avion,airplane*; *assiette,plate*; *jambon,ham*; *chemise, shirt*

During the next phase of the experiment, all 25-word pairs were presented and tested again, using a new grouping of words in the mixed-order acquisition condition but the same five-word groups in the grouped-order acquisition condition. These five-word groups and the five words pairs within each group were also presented in a new order. Three different groupings and orderings of word pairs were used. The sequence of these groupings and orderings was counterbalanced across participants. After the participants correctly responded to all 25 words on two consecutive trials, they saw all 25 French words one at a time and were asked to type in the corresponding English word (a new grouping of word pairs was used for the mixed-order acquisition condition and a new ordering of groups and of words within groups was used for the grouped-order acquisition condition).

One week later, participants returned and took the test again. Participants then repeated the acquisition phase to assess ease of relearning. The researchers had two different dependent measures: a) the number of errors (out of 25 possible) and b) total response time. They found that participants in the grouped-order acquisition condition made fewer errors than those in the mixed-order acquisition condition on the first trial. However, the participants in the mixed-order acquisition condition later made fewer errors than those in the grouped-order acquisition condition. The researchers also found that participants in the grouped-order acquisition condition were faster than those in the mixed-order acquisition condition during learning, but the opposite was found during relearning. In addition, participants were faster overall to respond during relearning than during learning.

Experiment 2

In the first experiment, participants continued with training until they reached a criterion of two consecutive trials with all responses correct. This method ensured that all participants reached the same level of performance at the end of training but led to the participants receiving different numbers of training trials. In order to overcome this problem in the second experiment, the participants (sixty non-French-speaking college students) were trained for a fixed number (three) of trials that did not depend on the participants' performance. The procedure and materials were the same as in the first experiment. Participants in the grouped-order acquisition condition made fewer errors than those in the mixed-order acquisition condition during learning, but the opposite was found during relearning. Again participants in the mixed-order acquisition condition made fewer errors than those in the grouped-order acquisition condition.

To summarise, Schneider, Healy and Bourne (1998) found that grouping lexical items according to topic-related vocabulary facilitated initial acquisition, but either hindered or had no effect on retention. In the first experiment during both learning and relearning, participants in the grouped initial acquisition condition responded more accurately than did those in the mixed initial acquisition condition on the first two trials but less accurately than did those in the mixed initial acquisition condition on the third trial. Furthermore, there was a trend toward faster responding on the part of participants in the grouped initial acquisition condition during the learning session but slower responding by the same group during the relearning session. In the second experiment, participants in the grouped initial acquisition condition made fewer errors than did those in the mixed initial acquisition condition during the learning session but made slightly more errors than did those in the mixed initial acquisition condition during the relearning session.

4.3.6 Finkbeiner and Nicol's study (2003)

In a more recent study (Finkbeiner and Nicol, 2003), participants (forty-seven undergraduates participating for course credit) learned 32 new L2 labels for concepts from four different semantic categories ('animals', 'kitchen utensils', 'furniture', 'body

part') in either a related or unrelated condition. During the test phase, participants were required to retrieve those labels in a translation task. The researchers manipulated semantic grouping both during training and during test. The test phase included both L1-L2 translation and L2-L1 translation. Thirty-two novel words[28] were created and each was paired with a picture of a familiar concept. Participants were seated individually in sound-resistant computer booths. During training, participants first heard a recording of the L2 word over headphones, then saw the L2 word and its corresponding picture for 500 ms on the monitor, and then heard a second recording of the L2 word. Participants were asked to repeat the L2 word twice into a microphone. The purpose of the repetitions was simply to facilitate learning.

In the 'related' training condition, semantically related items were blocked into groups of eight. Each block of eight was presented four times during training. In the 'unrelated' training condition, the 32 items were scrambled within a block and each block was presented four times. The vocabulary training was followed by a recognition task, which consisted of the presentation of a picture followed by one of the L2 labels. Participants were instructed to press a 'yes' button if the picture and the L2 word matched and a 'no' button if they did not. After the recognition task, participants were given the translation task (for example, in the L1-L2 blocks, an English word appeared and participants were asked to speak the 'L2' translation equivalent into the microphone as quickly as possible. The results revealed that participants translated L2 labels learned in semantic sets significantly more slowly than they did L2 labels learned in random order. This was the case in both translation directions. The results of this study suggest that both L1-L2 translation and L2-L1 translation involve semantic representations: there is no indication that L2-L1 translation bypasses the conceptual store (Finkbeiner and Nicol, 2003:378). This contradicts the Revised Hierarchical Model (RHM – see section 1.5.1) which proposes that, while forward translation is conceptually mediated, backward translation is lexically mediated (i.e. L2 words are directly linked to L1 words).

---

[28] The novel words were created to conform to English phonotactic constraints in order to reduce memory load. Half the words for each category were one syllable in length (e.g. *birk*) and the other half were two syllables in length (e.g. *valloon*).

4.4 Conclusion

The main purpose of the present chapter has been to examine two contrasting views of presenting and learning vocabulary. The arguments for the presentation of related lexical items together in sets are mainly based on theoretical rather than experimental evidence. Words can be related and grouped in various ways. This type of word grouping is called clustering. We examined the distinction between linguistically-based clustering (e.g. words grouped in lexical sets/semantic clusters or words grouped by sense relations), and cognitively-based clustering (e.g. thematic clustering).

There are three main arguments for the presentation of vocabulary in semantically related sets:

1) The first argument is that the presentation of semantically related vocabulary makes the meaning of these words clearer by seeing how they relate to and are different from other words in the set. Channell (1981) mentions that we should teach L2 vocabulary in related sets because the vocabulary of a language consists of interrelating networks of relations between words. We are no longer dealing with random lists of words, but with a systematic structure. This is important because, according to Channell (1981), the mind uses semantic similarity in classifying words.

2) The second argument is that there is evidence for the usability and effectiveness of presenting related vocabulary in classroom activities. Jullian (2000) refers to a classroom activity which incorporates an explicit approach towards the presentation of semantically related vocabulary. The students collect related words (close in meaning) for a given leading word (e.g. *hit*). They are asked to draw a semantic network around this particular leading word and illustrate the differences in meaning between the leading word (*hit*) and the related words (e.g. *strike, beat*). Jullian (2000) points out that this type of classroom activity helps students to understand the full semantic content of the related words and detect what makes them similar and different from each other. This classroom activity provides a useful framework for the students to understand and see where meaning overlaps. This is very important for vocabulary teaching because to understand a word fully and use it appropriately "what the learner needs to

know is how it relates to others of similar meaning and which other words it can be used with" (Carter and McCarthy, 1988:49).

3) The emphasis on the ability to distinguish differences between words with related meaning is also present in EFL coursebooks. Judging by a recent survey of ESL textbooks, it appears that many if not most ESL/EFL students are exposed to their new language vocabulary preorganized for them in semantic clusters (topic-related vocabulary). It appears that semantic clusters fit quite nicely into most current ESL textbooks. For example, learners are asked to learn 'foods' in *Headstart Beginner* Unit 5 (Beaven, 1995). There are also some books which exploit synonyms (*abandon/desert*) and antonyms (*hot/cold*). For example, *Gold proficiency* (Newbrook and Wilson, 2000:129) draws students' attention to the difference in meaning between *heal*, *treat* and *cure* by providing their dictionary definitions. Coursebook-writers are driven to present semantically related vocabulary items mostly because of their own perceptions of the communicative needs of their students. As a consequence, these coursebooks are divided into various units responding to any situation in which students might find it necessary to communicate in their language (e.g. visiting a doctor). Another reason why coursebook-writers present topic-related vocabulary is that semantic clusters fit naturally into 'slots' left open in the oral and written substitution activities. *New horizons in English 1* (Walker, 1991:41) offers *carrots, nuts, grapes, pears, peaches, oranges*, and three more food labels as possible ways to complete the question, 'Do you like _____ ?'

However, there is some experimental evidence against the presentation of semantically related vocabulary in sets. Tinkham (1993, 1997) and Waring (1997) investigated interference effects for word learning. In their studies, subjects listened to lists of English words paired with imaginary words. The English half of each word pair was presented and the participants had to remember and say the imaginary (L2) half of the pair within a set time. Their task was to learn (recall/retrieve) the meaning of the L2 words. The data collected by these researchers suggest that the presentation of new vocabulary items to L2 learners in clusters of semantically and syntactically similar words (*peach*, *apple*, *orange*) impedes rather than facilitates learning. Waring's (1997) and Tinkham's (1993) studies conclude that presenting students with wordlists of new

words in semantic clusters, rather than in unrelated word groups, can interfere with learning. This means that that it takes students more time to learn new lexical items when these lexical items are presented in related sets rather than presented in unrelated sets. The researchers found that it took from 47% to 97% more repetitions to learn the groups of related items, as compared to the number of repetitions it took to learn the group of unrelated items (Nation, 2000).

Nevertheless, there are some limitations (discussed by Waring, 1997) on the generalizability of the results found in the above studies. There are four main points we have to consider: a) very few words were tested, b) the artificial words that were used were strictly controlled, c) there is no clear definition of what semantic relatedness might mean and d) it is not clear whether these same effects will hold for learners who already have part of the semantic set being tested

It is also important to mention that contrary to the above researchers, Schneider, Healy and Bourne (1998) used natural-L2 words and not artificial words. The findings initially appeared to suggest that learning related words together (e.g. parts of the body) was easier than learning unrelated words. However, when a test of long-term retention was administered the results did not appear to be the same. The researchers found that the participants in the mixed-order acquisition condition (presented with unrelated vocabulary) were faster and made fewer errors than those in the grouped-order acquisition condition (presented with related vocabulary).

In a more recent study, Finkbeiner and Nicol (2003) used thirty-two novel words and each was paired with a picture of a familiar concept. The results revealed that participants translated L2 labels learned in semantic sets significantly more slowly than they did L2 labels learned in random order. This was the case in both translation directions. The results of this study suggest that both L1-L2 translation and L2-L1 translation involve semantic representations.

The arguments for and against the presentation in lexical sets reported in this chapter suggest further research. We have two contrasting views on the presentation of vocabulary in a L2 (experimental evidence vs. theoretical framework). However, we do

not have enough convincing evidence to decide which of the two contrasting approaches to vocabulary is the more useful and appropriate for vocabulary teaching in a L2. The best way for us to make a decision is to apply both approaches in EFL classrooms and compare the results. The following chapter describes a research plan that applies the two contrasting approaches to vocabulary teaching in English as a L2.

CHAPTER FIVE

Research Design

5.1 Introduction

The present chapter provides a detailed account of the structure of the research. There is a short description of the background to the research followed by the research paradigm we were inspired by. The rest of the chapter is divided into different sections which describe the research procedures and steps of the study. Chapter Five reports: the motivation for using Greek EFL students, the research aims, questions and hypothesis and the research procedure of the study. The selection of subjects, the selection of words, the teaching procedure and the tests used to test vocabulary knowledge are also reported in detail in the following sections.

5.2 Background to the research

In Chapter Four we discussed two opposing views of the use of semantically related vocabulary in L2 vocabulary teaching and learning. There is some experimental evidence which suggests that learning semantically related words (e.g. body parts) at the same time makes learning more difficult (Tinkham, 1993, 1997, Waring, 1997, Schneider, Healy and Bourne, 1998 and, Finkbeiner and Nicol, 2003). We also noted that there is a theoretical framework that strongly supports the idea that it is very useful to present words of related meaning together so that learners can see the distinctions between them and gain a complete coverage of the defined area of meaning (Channell, 1981, 1990).

The present research is based on the fact that we do not have a clear picture in which circumstances (i.e. for learners at different levels) the two approaches are helpful and useful to EFL students. The following paradox appears: while the experimental evidence suggests that semantically related vocabulary **does not** help vocabulary learning, the EFL coursebook-writers (based on the theoretical framework discussed in the previous chapter) present vocabulary in semantic clusters (topic-related vocabulary). The experimental evidence mainly derives from research using artificial language and not a

natural L2 (Tinkham, 1993, 1997, Waring, 1997). Although Schneider, Healy and Bourne (1998) used French as a L2, their experiment was not applied to natural foreign language students in a L2 classroom. The purpose of our research is to investigate which of the two contrasting views (described earlier) will prove to be a useful tool in L2 vocabulary learning. We want to throw some new light on this topic (in other words to carry out *illuminative research* according to Wallace, 1998:43) and discover whether the findings of previous research actually apply to a natural teaching environment using a real L2. The best way to do that is to conduct our study in a natural EFL classroom. If we want to enrich our understanding of language learning and teaching, we need to spend time looking in classrooms with natural learners. For this reason, we will follow a plan that is partly action-research oriented. It is not entirely action-research-based because we plan to test students' vocabulary learning. In other words, we are using elements of action research combined with elements of quasi-experimental design. The next sections give a detailed account of the key features of action research and the elements of quasi-experimental design used in the present study.

5.3 Action-research orientation with elements of quasi-experimental design

A form of research which is becoming increasingly significant in language education is action research. Cohen and Manion (1994:186) define action research as a "small-scale intervention in the functioning of the real world and a close examination of the effects of such an intervention". Action research is normally associated with small-scale research projects. It is designed to bridge the gap between research and practice (Somekh, 1995:340). It combines diagnosis with reflection, focusing on practical issues that have been identified by participants and which are somehow both problematic yet capable of being changed (Elliott, 1991:49). Action research is problem-focused in its approach and very practical in its intended outcomes. According to Wallace (1998:4), action research is a way of reflecting on your teaching and is done by systematically collecting data on our everyday practice and analysing it in order to come to some decisions about what our future practice should be. In this sense, action research is located within the context of strategies for professional (i.e. teaching) development.

Denscombe (1998:58) points out that action research is concerned with the aims of research, but does not specify any constraints when it comes to the means for data collection that might be adopted by the action researcher. This leaves room to use elements of quasi-experiment since we plan to test students' vocabulary learning by using vocabulary tests. These tests require quantitative analysis by means of statistical techniques and statistical tests of significance (that give researchers additional credibility in terms of the interpretation they make). These kinds of statistical techniques and tests are mainly used in true or quasi-experiments to provide solid foundations for description and analysis. In other words, the present study has elements of quasi-experimental design in terms of data collection and analysis.

In this sense, our study is action research-oriented to a certain extent. It is inspired by an action-research paradigm and it is not a true experiment or even a pure quasi-experiment. It is preferred to the experimental techniques as true experiments are relatively rare in education because the researcher often has little choice in the matter of subjects and therefore has to work with existing groups. The researchers generally do not have full control over the independent variables and are unable to randomly allocate subjects to different treatment conditions. True experiments involve at least one control group (which has been or will be exposed to the independent variable) and one experimental group both randomly constituted. According to Nunan (1992:41), a true experiment has both pre- and post-tests, experimental and control groups and, random assignment of subjects. On the other hand, a quasi-experiment has both pre- and post-tests and experimental and control groups, but no random assignment of subjects. Its essential feature is a lack of random allocation to different treatment conditions.

At this point it is useful to mention that action research has common elements with case study. It is mentioned that action research frequently uses case study (Cohen, Manion and Morrison, 2000:181). Case study is an approach which fits very comfortably into the action research framework because it is tightly focused and personalized (Wallace, 1998:170). It provides a unique example of real people in real situations (e.g. a particular class or school). It is a small-scale research which focuses on particular

individuals or groups and provides a chronological narrative of events relevant to the case (Hitchcock and Hughes, 1995:322)[29].

One of the strengths of the case study approach (similar to the action research approach) is that it allows the researcher to use a variety of sources, methods and types of data (both *qualitative* i.e. field notes, interviews and *quantitative* i.e. numeric-statistical data) as part of the investigation (Denscombe, 1998:31).

According to Cohen, Manion and Morrison (2000:226), action research can be used in a variety of areas. For example in:

1) teaching methods by replacing a traditional method (i.e. the grammar-translation or traditional method) by a discovery method (i.e. direct method, audiolingual method, etc.);
2) learning strategies by adopting an integrated approach to learning (teaching of multiple subjects at the same time) in preference to a single-subject style of teaching and learning;
3) evaluative procedures by improving one's methods of continuous assessment;

In this sense, action research is an approach to improving education by changing it and learning from the consequences appearing in classroom and schools. Action research is concerned with action, with doing things and translates generalizations into specific acts. It is both 'action' and 'research' (Cohen, Manion and Morrison, 2000:227). It bridges the gap between practical theories and practice using instruments of data collection such as questionnaires, interviews and observational data (Cohen, Manion and Morrison, 2000:237).

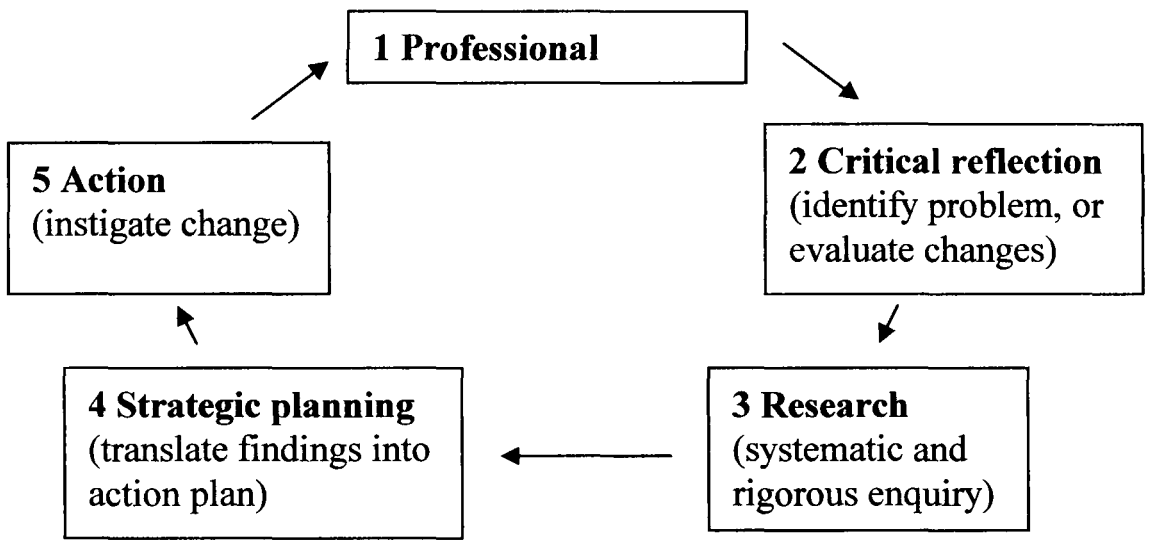
5.3.1 Four defining characteristics of action research

According to Denscombe (1998:57-58), there are four defining characteristics of action research:

---

[29] In Cohen and Manion and Morrison (2000:182).

1) ***Practical***. Action research is aimed at dealing with real-world problems and issues, typically at work and in organizational settings. It is driven by the need to solve practical, real-world problems. It encourages the practitioner to investigate his or her own practices with a view to altering these in a beneficial way.
2) ***Change***. Change is regarded as an integral part of research. Action research is wedded to the idea that change is good. Because, action research tends to be localized and small-scale, it usually focuses on change at the micro level.
3) ***Cyclical process***. The crucial points about the cycle of inquiry in action research are (a) that research feeds back directly into practice, and (b) that the process is ongoing. Initial findings generate possibilities for change which are then implemented and evaluated as a prelude to further investigation. The cyclical process in action research according to Denscombe (1998) looks like this:



Denscombe (1998:60)

4) ***Participation***. Action research involves the practitioners very closely. Practitioners (for our purposes me as a researcher for my doctoral work) are the crucial people in the research process. Their participation is active, not passive. The participatory nature of action research is probably its most distinctive feature. Action research insists that practitioners must be participants, not just in the sense of taking part in the research but in the sense of being a partner in the research.

5.3.2 Advantages of action research

According to Denscombe (1998), there are four main advantages of action research:

1) It addresses practical problems in a positive way, feeding the results of research directly back into practice.
2) It has personal benefits for the practitioner, as it contributes to professional self-development.
3) It should entail a continuous cycle of development and change via on-site research in the workplace, which has benefits for the organization to the extent that it is geared to improving practice and resolving problems.
4) It involves participation in the research for practitioners. This can democratize the research process, depending on the nature of the partnership, and generally involves a greater appreciation of, and respect for, practitioner knowledge.

Denscombe (1998:65)

Action research, then, leads to equality and co-operation and produces practical solutions to isolated problems. In this sense, it is emancipatory. Cohen, Manion and Morrison (2000:232) mention that the emancipatory interest of action research is based on the notion of action researchers as participants in a community of equals. Action research involves the significance of personal participation and improves professional action by investigating professional practice.

5.3.3 Disadvantages of action research

Cohen, Manion and Morrison (2000:233) based on several critical studies that appear in the literature provide a list of several drawbacks and negative aspects of action research. It is suggested that action research:

1) is utopian and unrealizable. Denscombe (1998:65-66) mentions that the nature of the research is constrained by what is permissible and ethical within the workplace setting
2) is too controlling and prescriptive. The setting for the research does not allow for the variables to be manipulated, because the research is conducted not alongside routine activity but actually as part of that activity (Denscombe, 1998:65-66)
3) is uncritical and self-contradicting in the sense that the action researcher is unlikely to be detached and impartial in his or her approach to the research
4) is naive in its understanding of groups and celebrates groups over individuals

I also believe that care needs to be taken over the reliability and validity of all forms of research including action research. Caution should be exercised in making strong claims about the research outcomes. The involvement of the practitioner limits the scope and scale of research. Denscombe (1998:65) points out that the 'work-site' approach (of action research) affects the representativeness of the findings and the extent to which generalizations can be made on the basis of the results. Action research also tends to involve an extra burden of work for the practitioners. Nunan (1992:18) mentions that collaboration should not be seen as a defining characteristic of action research because many teachers are either unable, for practical reasons, or unwilling, for personal reasons, to do collaborative research.

5.3.4 Process of action research

Cohen, Manion and Morrison (2000:235-236) propose an eight-stage process of action research which is summarized below. Action research is:

- initiated by a question (derived from a real problem in the classroom which needs to be confronted),
- reviews the literature to find out what can be learned from comparable studies (their objectives, procedures and problems encountered),
- is supported by objective data and interpretation (conditions and methods of data collection, monitoring of tasks, classification and analysis of data), and
- is carried out by a practitioner investigating aspects of his or her own context and situation.

Finally, the project takes the form of an ongoing cycle (see section 5.3.1) in which the practitioner reflects on, returns to, and extends the initial inquiry. This final stage includes discussion of the findings, consideration of any mistakes and problems found in the project and recommendations for further research.

5.4 Motivation for using Greek EFL students in the research

Our motivation for using Greek EFL students as subjects in our study derives from: a) myself having been a learner of English as a L2, since I am a Greek native speaker receiving higher education in the U.K., b) my personal experience in teaching English as

a L2 in private schools in Greece and, c) a recent paper by Scholfield and Gitsaki (1996). For their study, the researchers used learners' reports about how they are taught and how they learn new English vocabulary as a guide to examining the difference between the two prevailing pedagogical environments in Greece: private institutions of foreign languages ('frontistiria') and public (state) schools. One hundred and eighty elementary or pre-intermediate learners participated in their survey using questionnaires including closed and open-ended questions. There are four interesting findings to emerge from Scholfield and Gitsakis' (1996) study:

1) that writing new words and their translations in Greek was the commonest kind of note kept on new vocabulary,
2) new vocabulary is seen as a simple bilingual rather than monolingual list of words,
3) English synonyms are rarely used and,
4) studying vocabulary is a process done predominantly at home by students because teachers do not devote much time to vocabulary teaching.

Scholfield and Gitsaki (1996) expect students at more advanced stages of learning to be more sophisticated than those at lower levels in the process of learning new vocabulary. They expect them: (a) to be able to distinguish in more detail between different words with similar meaning; (b) to keep fuller notes about the new words they learn; (c) to classify new words according to their meaning or their collocation; (d) to get English explanations more frequently as their level of English would be sufficient for this kind of practice.

Scholfield and Gitsaki (1996) also claim that poor vocabulary knowledge and repetitious writing characterize the majority of Greek EFL students. They overuse the most frequent words providing a flat and sometimes uninteresting style in their essays. Occasionally, Greek EFL students even fail to express the ideas they want to communicate. Scholfield and Gitsaki (1996:125) argue that synonymy and antonymy exercises could be one way of introducing the meaning of new words.

There are several points from Scholfield and Gitsakis' (1996) study that inspired the present researcher to use Greek EFL students as subjects. The main point is that most of

the Greek students seem to have relatively poor vocabulary knowledge and tend to use a small number of words in their writing. We want to examine if learning semantically related (topic-related vocabulary, synonyms, antonyms and homonyms) or unrelated words will help students *learn*[30] and *recall*[31] more vocabulary items. Concerning the point that new vocabulary is seen as a simple bilingual rather than monolingual list, it is worth examining whether it can be a beneficial or confusing factor in L2 vocabulary acquisition. Also given that English synonyms are rarely used, it is interesting to see if synonymy and antonymy exercises could be one way of introducing the meaning of new words.

5.5 Overview of research

We will conduct a study of alternative ways of teaching vocabulary to Greek EFL students. There will be a classroom study which will employ two different ways of organising new vocabulary for presentation: a) presenting semantically related words (words that share certain semantic and syntactic similarities, for example, topic-related vocabulary i.e. *knife, fork, spoon*, synonyms, antonyms or homonyms) together at the same time, and b) presenting vocabulary in an unrelated (mixed) fashion (words that are not semantically related i.e. *book, hospital, freedom*). The aim of the present study is to evaluate the relative claims of the two different procedures by using two different groups of students (Class A and Class B). At the end of the research period, all students will be tested to determine which of the two competing methods is the more effective (the next sections of this chapter provide a detailed account of the present research procedure).

5.6 Research questions and hypotheses

There are four main questions and other subsidiary questions to be examined in the present study. The first is concerned with the way the words are taught (presented), the second refers to whether word properties (analysed in section 5.7.2.2) affect scores on

---

[30] The written form and the semantic meaning of a L2 word.

[31] Be able to elicit the target word from memory when they are provided with some stimulus. A recall item requires the test-taker to provide the required form or meaning (Nation, 2001:359).

vocabulary tests and the third has to do with the type of vocabulary presentation intermediate students prefer. Let me now analyse each question separately.

The **first research question** is as follows: Which of two ways of presenting and organizing the teaching of new L2 vocabulary (specifically related word sets or unrelated word sets) produces better retention of those words when retention is operationalized as scores in a short-term and a long-term vocabulary translation test? There are two informal hypotheses to be tested:

Hypothesis 1: When tested on their knowledge of 60 new L2 words (after the teaching period of three weeks), subjects will achieve a higher test score when tested on related vocabulary compared to their test score on their knowledge of 60 unrelated words.

Alternative Hypothesis : When tested on their knowledge of 60 new L2 words (after the teaching period of three weeks), subjects will achieve a higher test score when tested on unrelated vocabulary compared to their test score on their knowledge of 60 related words.

The first research question will be answered using t-tests.

The **second research question** is as follows: Which properties of L2 words (depending on the manner of their presentation in either related or unrelated word sets) seem to facilitate retention (short-term and long-term) when retention is operationalized as scores in a short-term and a long-term vocabulary translation test? The properties of L2 words that will be examined (in relation to the manner of their presentation) are: word frequency, word length and word concreteness/abstractness. The second research question will be answered using ANOVA.

The **third research question** is as follows: Which properties of L2 words (independent of the manner of their presentation in either related or unrelated word sets) seem to facilitate retention (short-term and long-term) when retention is operationalized as scores in a short-term and a long-term vocabulary translation test? The frequency of a word being important for learning is a factor to be examined in this study. We want to enquire if frequency as measured by the BNC predicts ease of learning as measured by scores on vocabulary tests. This part will be examined by using Spearman Rank Order Correlation. How are these scores also affected by word length, concreteness and

abstractness? Which factor is more powerful than others? The third research question will be answered using Factorial ANOVA.

The **fourth research question** is as follows: Which methods of presenting new L2 vocabulary do intermediate subjects prefer? This question will be answered using questionnaires.

Subsidiary questions of the research are as follows: Is the effect of unrelated vs related vocabulary (on test scores) influenced or mediated by subjects' level (beginners vs intermediate) and age (children vs adults)? Is there a difference between test scores between short-term tests and long-term tests (SHT tests vs LT tests)? Is there a difference in test scores between Class A and Class B for children and adults separately (Class A vs Class B)? All these questions will be answered using t-tests. In addition, which of the semantically related word sets (antonyms, synonyms, homonyms and topic-related) facilitates or hinders subjects' performance in vocabulary test scores? This question will be answered using ANOVA.

5.7 Research procedure

The research procedure of this study consists of four main steps:

a. selection of subjects
b. selection of words
c. teaching procedure
d. testing vocabulary knowledge and ways of testing

Now I want to discuss each section (and sub-sections) of the research procedure separately.

5.7.1 Selection of subjects

In designing studies we would like to select our subjects randomly from the largest population possible in order to get results that have the greatest generalisability. As Hinton (2004) points out, researchers inevitably compromise, however, and lose some generalisability in favour of greater control over the variables involved.

The sample used in this study is a non-probability sample. This means that every member of the wider population does not have an equal chance of being included in the sample. A non-probability sample derives from the researcher targeting a particular group, in the full knowledge that it does not represent the wider population (Cohen, Manion and Morrison, 2000:102). In other words, it simply represents itself. This type of sample is used in small-scale research, for example, two or three groups of students, where no attempt to generalize is desired (Cohen, Manion and Morrison, 2000:102). Thus, non-probability samples are frequently used in action research studies. There are different types of non-probability samples. The sample used in this study can be characterized as a convenience sample (or as it is sometimes called, accidental or opportunity sampling), and a purposive sample. Convenience sampling involves choosing the nearest individuals to serve as subjects (the researchers simply chooses the sample from those to whom she has easy access), and purposive sampling involves choosing a sample that is satisfactory to the researcher's needs (Cohen, Manion and Morrison, 2000:103). In both cases, the sample does not seek to generalize about the wider population.

The subjects in this study are classified in two groups, intermediate EFL children and beginners EFL adults. This sample is first characterized as a convenience sample because it is the only one I had access to and secondly as useful in terms of examining the learnability effects of the two different levels. Taking into consideration the difficulties and complexities we faced to get permission to conduct our study in a real EFL classroom environment (due to limited resources, time and access to subjects), we were fortunate to find two educational institutions that allowed us to enter EFL classroom and teach new English vocabulary. The fact that we had intermediate and beginners EFL learners as subjects in our study provided us with the opportunity to answer the question regarding the learning process of intermediate learners proposed by Waring (1997:269). According to him, the findings of his study applied only to beginning EFL learners. Thus, it would be interesting to see how using semantically related and unrelated vocabulary would affect the performance of both beginning and intermediate EFL learners.

The same study was conducted twice in two different time periods: in the first case we obtained official permission to undertake the research on 24/2/2004, while in the second case access was only given on 22/10/2004, eight months later (see next sections for further details and explanations of each part of the study). The two groups mentioned above are the only samples I was able to test. This sample might not generalize to all L2 learners of the English language, yet this should not stop the researcher carrying out the research because important information can still be found. Now I want to talk about each group (sample) separately.

The first sample consists of 31 Greek EFL young learners (20 girls and 11 boys), intermediate level, and aged from 11 to 16 years old, who are studying for the First Certificate (FC) Cambridge examination[32] in a private educational institution in Greece. The students have been studying English for five or six years, and all have both parents Greek-speaking. Two intermediate existing classes (Class A with 16 students and Class B with 15 students) participated in our study. Most of the students learn a second foreign language in the same private school, 15 students learn German and 7 students learn French. All students study English at the Greek (state) school as well. They spend three to four hours per week studying English. The students' favourite skill is *speaking* while *writing* seems to be the most difficult. All the information mentioned above was obtained from the Students' Language Background Questionnaire (see section 5.7.3 and Appendix 4). Greek students attend private classes in the afternoon. The Greek EFL students who took part in this study attended classes every Monday, Wednesday and Friday. The duration of each lesson every Monday and Friday was an hour and a half (Class A from 16:30 to 18:00 pm and Class B from 18:00 to 19:30 pm) while the lesson every Wednesday lasted only forty-five minutes (Class A from 16:30 to 17:15 pm and Class B from 17:15 to 18:00 pm). In order to have enough time to present and teach new vocabulary to the students, we decided to teach the classes every Monday and Friday. All students were taught English using the same coursebook called *Shine C* (Garton-Sprenger, J. and Prowse, P., 2000). *Shine C* is intended for intermediate students and consists of nine units on different topics divided into four lessons. Each lesson corresponds to activities and tasks encouraging the four skills (speaking, reading, listening and writing).

---

[32] The students will sit the exams in two years time.

The second sample consists of Greek EFL adult learners. We found two groups of adult beginners. They are public servants who attend English classes (seminars) to improve their basic knowledge of the English language. The seminars are sponsored (and financed) by the Greek government and their purpose is to develop the educational skills of public servants. These seminars lasted from September 6th 2004 to December 31st 2004. The courses are designed and addressed to all public servants who are interested in attending the seminars in order to improve their writing and speaking skills in English. At the end of the courses, each participant receives a certificate of attendance.

The subjects in the second sample are 32 Greek adult beginners (22 female and 10 male), aged from 30 to 50 years old, who attend adult-classes (seminars) on English language in Greece. The participants have been studying English for six weeks. Two existing classes (Class A with 17 students and Class B with 15 students) participated in our study. Only three participants are learning another second language; two participants are learning German and one French (information obtained from the Participants' Language Background Questionnaire - see section 5.7.3 and Appendix 10). Both groups are taught English by the same teacher. Most of the participants have children who study English as their second language at Greek public school. The adults attend the classes for personal and professional needs, and their favourite skill is *listening* while *speaking* is the most difficult for them. The participants attend classes in the evening. The adults who took part in this study attended classes twice a week; Class A on Mondays and Fridays (17:00 – 18:30 pm), and Class B on Mondays and Fridays (19:00 – 20:30 pm). The duration of each lesson was an hour and a half. The course was based on material prepared by their teacher (photocopies with activities and tasks encouraging the four skills: speaking, reading, listening and writing).

5.7.2 Sources for selecting vocabulary to be taught

The vocabulary items used for this study were selected from two different sources: a) the *English Vocabulary in Use* (McCarthy and O'Dell, 2001) and b) the *Longman Vocabulary Website* (www.ablongman.com/vocabulary). We chose these particular sources because they apply to intermediate students (like our first sample) and help them to expand their vocabulary knowledge to upper-intermediate level. Even though the

students already use an intermediate coursebook, the specific sources were selected in order to control the quality and quantity of vocabulary items used in the study. Let us examine each source separately.

a) The *English Vocabulary in Use* is an upper-intermediate EFL vocabulary coursebook. This coursebook is written to help EFL students (intermediate level) improve their English vocabulary. The selection of words appearing in the coursebook is based on the frequency data from the Cambridge International Corpus to ensure that words are not too advanced or too easy for this particular level. The book is divided into a number of sections. It has 100 two-page units. In most units, the left-hand page explains the words and expressions to be studied in that unit and the right-hand page provides a series of exercises practising what the students have just learned. The coursebook presents vocabulary items organized in topic-related units (e.g. sport, music, food, etc).

b) *The Longman Vocabulary Website* (www.ablongman.com/vocabulary) is a vocabulary website for student resources. It provides a list of vocabulary categories such as *word parts, synonyms, antonyms, homonyms, easily confused words,* etc. Students can double-click on the category that interests them or that they need extra practice in. The students are then taken to many exercises that will help them strengthen and increase their vocabulary knowledge on that category. The exercises are grouped into beginner, intermediate, and advanced levels. We selected words from all three levels to obtain a larger variety of synonyms, antonyms and homonyms.

5.7.2.1 Principles for selecting vocabulary to be taught

It is knowledge of content words (nouns, 'full' verbs, adjectives and adverbs) that we focus on. In order to be as concise and consistent as possible, we will present words which belong to the same lexical category. We will present students with nouns because, according to Ellis and Beaton (1993b), nouns are easier to learn as learners can form mental images of them more readily. It is important to point out that a noun could simultaneously meet more than one criterion. There are seven basic criteria for selecting nouns. A noun could be:

i) concrete or abstract noun

There is a distinction between concrete and abstract nouns. Concrete nouns name things that can be seen or touched. By contrast, an abstract noun represents an idea, experience or quality rather than an object that can be touched. For example, *sadness* is an abstract noun but *table* is not. However, there are some nouns that are not clearly classified as abstract or concrete. For example, *drought*, *flood* and *murder* are not clearly categorized as concrete or abstract. We will treat these nouns as concrete nouns because they name something that can be seen in the real world and can easily be represented in a picture.

ii) countable, uncountable or collective noun

Countable nouns can be counted (e.g. *book*) while uncountable nouns cannot (e.g. *air*). In addition, collective nouns can take a singular form but are composed of more than one individual person or items (e.g. *team*).

iii) non-compound noun

Compounding is the process of forming a word by combining two or more existing words. For example, words like *bedroom* and *newspaper* are excluded from our list.

iv) singular noun

v) not cognate

These are words related by derivation, borrowing, or descent. In this case, we exclude all English nouns that derive from Greek (e.g. *dogma, oregano*). All nouns must originate from middle/old English, French, German or Latin.

vi) short (monosyllabic) or long (polysyllabic) noun

Intuitively, it would seem that longer words should be more difficult simply because there is more to learn and remember. Yet the empirical results are not conclusive. Laufer (1997:145) states that one common misperception is to assume that shorter words are easier because they are more frequent in language (e.g. in English). This does not mean that in other languages short and frequent necessarily go hand-in-hand. For example, a large component of Greek vocabulary consists of long and frequent words (i.e. *πραγματικότητα,* which means 'reality', belongs to the first 1000 most frequent words in Greek language, see HNC, http://hnc.ilsp.gr/en/info.htm). Thus, in a learning situation it is hard to attribute the difficulty of recalling a particular word to its length rather than to a variety of factors. What can account for better learnability is not the word's length, but any of the learner's frequent exposure to it. Laufer (1997) points out that it is the quantity of input that may contribute to the successful learning of the short words, not

their intrinsic quality. In this study, short words are meant to have one syllable (e.g. *ebb*), while long words have two or more syllables (e.g. *blister*, *estuary*).

vii) frequent or non-frequent

Material written for EFL students needs to use simplified vocabulary and structure if it is to be accessible to lower and intermediate level students. Textbook writers need to keep vocabulary simple. Some words are more frequent in use than others. The General Service List (GSL) (West, 1953) is the specific list of 2,000 words that Nation (1990) refers to when he writes about the 'first 2,000 words'. It is based on written texts, it is old, and it is not in frequency order, though frequency numbers are given. The list was compiled based on frequency alone. It was created to be an ideal vocabulary for EFL students to start out with. We used the General Service List to identify the frequency of the words used in this study.

5.7.2.2 Procedure to prepare the final list of words

The first step is to decide how many vocabulary items to teach in each lesson. Nation (2001:93) mentions that if too many words are focused on, they are likely to be forgotten or become confused with each other. He suggests we should deal with a few words in each lesson, probably 5 or 6 at the most. Schmitt (2000:144) also mentions that numbers in the area of ten words per 1-hour session are sometimes cited, and this does not seem unreasonable. For this reason, we decided to present ten words per lesson. This means that we had to collect sixty words for Class A (twenty words per week for a period of three weeks) and sixty words for Class B.

In order to have 120 words in total and have a sufficient number of words to choose from we created a list of more than 300 words (at least twice the number of words needed for this study). We have therefore produced a list of 306 randomly selected nouns. The list included topic-related vocabulary, homonyms, synonyms and antonyms. Once a sample of words has been selected, it is essential to find out whether the students already knew each word. One way is to ask students to supply a L1 equivalent for each L2 target word (Read, 2000). Another way is to use the checklist (or yes-no) test. In this test, target words are presented on a list and learners are required to check (tick) if they know them or not (Schmitt, 2000:174). The checklist test is really the simplest possible

format for testing vocabulary (Read, 2000:88). The problem is that many subjects might overestimate their vocabulary knowledge and tick words they really do not know. To overcome this problem, non-words[33] that look like real words but are not, such as *flinder*, are put into the test along with the real words (Schmitt, 2000:175). If some of the non-words are ticked, that indicates that the student is overestimating his or her vocabulary knowledge (see Anderson and Freebody, 1983, in Read, 1997:312). For this study, we used a checklist test. All the selected words were listed in alphabetical order. We then put some non-words (e.g. *dogner*, *glotune*) onto the list. We used 17 non-words in total (one non-word after 18 real words). So, the total number of words in the checklist was 323. The students are asked to tick the words they know. The checklist was administered to EFL children (both class A and B) on 27/02/2004. The students did not tick any non-words. Appendix 1 provides a sample of the L2 checklist test distributed to the intermediate children.

From the remaining words (words that EFL children had not ticked) we selected 120 words, 60 semantically related nouns and 60 semantically unrelated nouns (see Tables 5.1 and 5.2) The semantically related nouns were divided into six groups of ten words (three groups of topic-related nouns, one group of homonyms, one group of synonyms and one group of antonyms). The three topic-related groups were based on three different topics: a) crime, b) nature and c) food. The homonyms, synonyms and antonyms were organised in pairs of two (e.g. *pane-pane, torment-torture, ebb-flow*). The semantically unrelated nouns were also divided into six groups (ten words in each group). The words were randomly allocated in each group making sure that the words in each group are not semantically related. We have to mention here that the checklist mentioned above was also administered to the EFL adult groups on 22/10/2004. The students did not tick any non-words. They also ticked a very small number of the words used in the list (see section 5.7.3 and Appendix 11). For this reason, we were able to present and teach exactly the same words to both children and adult EFL groups. Tables 5.1 and 5.2 present the semantically related and unrelated nouns used in this study.

---

[33] Non-words are also used in the lexical decision task. This technique presents a sequence of letters on a computer display and asks the subject of the experiment to decide as quickly as possible whether it is a familiar word or not. Half of the items are actual (real) words, and half are not words at all, although they look like good potential words i.e. *flink* (Kenneth, Foster and Nan Jiang, in Nicol (ed), 2001:72).

Table 5.1: Semantically related nouns

Group 1-'Crime'	Group 2-'Nature'	Group 3-'Food'
Topic-related vocabulary	**Topic-related vocabulary**	**Topic-related vocabulary**
smuggling	cape	lamb
terrorism	peninsula	herring
forgery	cove	veal
mugging	tributary	ham
trial	valley	cod
proof	gorge	trout
jury	stream	prawn
verdict	estuary	shrimp
witness	ridge	squid
bribery	summit	lobster
Group 4	**Group 5**	**Group 6**
Homonyms	**Synonyms**	**Antonyms**
pane	torment	ebb
pain	torture	flow
steak	jab	gloom
stake	punch	glee
toe	spat	certitude
tow	quarrel	doubt
colonel	gleam	loyalty
kernel	twinkle	treason
council	boredom	poverty
counsel	tedium	prosperity

Table 5.2: Semantically unrelated nouns

Group 1	Group 2	Group 3
invasion	tube	rein
mortgage	tornado	loan
menace	sage	flare
controversy	carpenter	bane
custom	pigeon	fur
soul	bruise	plea
mussel	waist	blister
tailor	sensor	signet
excess	dough	landing
evidence	willow	ascent
Group 4	**Group 5**	**Group 6**
peril	bud	rash
query	plumber	tool
hoax	jug	jest
raid	whisker	quest
bias	porch	jeopardy
hatred	leek	scent
mane	peel	account
pollen	creek	disdain
famine	hail	prejudice
assent	tee	animosity

It was ensured that these nouns fulfil the criteria for selecting words discussed earlier. The description of the properties of the words used in this study is provided in Table 5.3. The words are in alphabetical order. The total number of words used in the study is 120. Table 5.3 includes 31 frequent words and 89 infrequent words (frequency is defined according to West's GSL as mentioned earlier), 57 short and 63 long words as well as 61 concrete and 59 abstract words. According to Laufer (1997), word length and word abstractness/concreteness are factors affecting word learnability (see section 2.2.1). Word frequency, on the other hand, has two aspects: a) the frequency in language and b) the exposure in the classroom environment. The relationship between word frequency in language and word learning is thought to be analogous.

The frequency of a word, being important for learning (according to theories and models of lexical profiles[34]), is a factor to be examined in this study. It is important though to make the following distinction. The theory states that the more frequent a word is in language then the more likely it is to be learned. Even though we notice most words in Table 5.3 are infrequent (according to our classification) it needs to be mentioned that word frequency in the English language may differ from word frequency in a classroom environment (or out-of-class environment). The exposure a student has to a certain word through teaching can be more or less effective regardless of the word's frequency in language. The frequency must be seen as a usage factor dependent on the type of language input that the learner receives. Laufer (1997:141) points out that "the frequency of a word's occurrence may be much different in a naturalistic, all-purpose language course as compared to a course in language for specific purposes". For the purpose of this study, word frequency will only be considered as frequency in language and not as a degree of word exposure in teaching.

[34] Theories and models of lexical profiles lie behind the use of frequency data of such things as learning lists (i.e. General Service List, Academic Word List, etc). An extensive list of theories and models can be found in Nation, 2001:9-21.

Table 5.3: Description of properties of words

	Frequent (First 2,000 words)	**Not-Frequent** (Academic Word List) and (Off-list words)	**Short** (one syllable words)	**Long** (two or more syllable words)	**Concrete**	**Abstract**
Number of words	31	89	57	63	61	59
account	√			√		√
animosity		√		√		√
ascent		√		√		√
assent		√		√		√
bane		√	√			√
bias		√		√		√
blister		√		√	√	
boredom		√		√		√
bribery	√			√		√
bruise		√	√		√	
bud		√	√		√	
cape	√		√		√	
carpenter		√		√	√	
certitude		√		√		√
cod		√	√		√	
colonel		√		√	√	
controversy		√		√		√
council	√			√		√
counsel		√		√		√
cove		√	√		√	
creek		√	√		√	
custom	√			√		√
disdain		√		√		√
doubt	√		√			√
dough		√	√		√	
ebb		√	√		√	
estuary		√		√	√	
evidence		√		√		√
excess	√			√		√
famine		√		√		√
flare		√	√		√	
flow	√		√		√	
forgery		√		√		√
fur	√		√		√	
gleam		√	√			√
glee		√	√			√
gloom		√	√			√
gorge		√	√		√	
hail		√	√		√	
ham		√	√		√	
hatred	√			√		√
herring		√		√	√	

hoax		√	√			√
invasion		√		√		√
jab		√	√			√
jeopardy		√		√		√
jest		√	√			√
jug		√	√		√	
jury		√		√		√
kernel		√		√	√	
lamb		√	√		√	
landing	√			√	√	
leek		√	√		√	
loan	√		√			√
lobster		√		√	√	
loyalty	√			√		√
mane		√	√		√	
menace		√		√		√
mortgage		√		√		√
mugging		√		√		√
mussel		√		√	√	
pain	√		√			√
pane		√	√		√	
peel		√	√		√	
peninsula		√		√	√	
peril		√		√		√
pigeon	√			√	√	
plea		√	√			√
plumber		√		√	√	
pollen		√		√	√	
porch		√	√		√	
poverty	√			√		√
prawn		√	√		√	
prejudice	√			√		√
proof	√		√			√
prosperity		√		√		√
punch		√	√			√
quarrel	√			√		√
query		√		√		√
quest		√	√			√
raid		√	√			√
rash		√	√		√	
rein		√	√		√	
ridge		√	√		√	
sage		√	√		√	
scent	√		√			√
sensor		√		√	√	
shrimp		√	√		√	
signet		√		√	√	
smuggling		√		√		√
soul	√		√			√
spat	√		√			√
squid		√	√		√	
stake		√	√		√	
steak		√	√		√	
stream	√		√		√	
summit		√		√	√	
tailor	√			√	√	
tedium		√		√		√

tee		√	√		√	
terrorism		√		√		√
toe	√		√		√	
tool	√		√		√	
torment		√		√		√
tornado		√		√	√	
torture		√		√		√
tow		√	√			√
treason		√		√		√
trial	√			√		√
tributary		√		√	√	
trout		√	√		√	
tube	√		√		√	
twinkle		√		√		√
valley	√			√	√	
veal		√	√		√	
verdict		√		√		√
waist	√		√		√	
whisker		√		√	√	
willow		√		√	√	
witness	√			√	√	

The next table (Table 5.4) presents a clearer picture of the properties of words in relation to the distinction of semantically related and unrelated vocabulary explained earlier.

Table 5.4: List of words used in the study (N=60)

Related (N=60)		**Unrelated (N=60)**	
Frequent (N=16)		**Frequent (N=15)**	
bribery	proof	account	prejudice
cape	quarrel	custom	scent
council	spat	excess	soul
doubt	stream	fur	tailor
flow	toe	hatred	tool
loyalty	trial	landing	tube
pain	valley	loan	waist
poverty	witness	pigeon	

Infrequent (N=44)		
boredom	jury	summit
certitude	kernel	tedium
cod	lamb	terrorism
colonel	lobster	torment
counsel	mugging	torture
cove	pane	tow
ebb	peninsula	treason
estuary	prawn	tributary
forgery	prosperity	trout
gleam	punch	twinkle
glee	ridge	veal
gloom	shrimp	verdict
gorge	smuggling	
ham	squid	
herring	stake	
jab	steak	

Infrequent (N=45)		
animosity	hail	porch
ascent	hoax	query
assent	invasion	quest
bane	jeopardy	raid
bias	jest	rash
blister	jug	rein
bruise	leek	sage
bud	mane	sensor
carpenter	menace	signet
controversy	mortgage	tee
creek	mussel	tornado
disdain	peel	whisker
dough	peril	willow
evidence	plea	
famine	plumber	
flare	pollen	

Short (N=29)		
cape	ham	spat
cod	jab	squid
cove	lamb	stake
doubt	pain	steak
ebb	pane	stream
flow	prawn	toe
gleam	proof	tow
glee	punch	trout
gloom	ridge	veal
gorge	shrimp	

Short (N=28)		
bane	jug	rein
bruise	leek	sage
bud	loan	scent
creek	mane	soul
dough	peel	tee
flare	plea	tool
fur	porch	tube
hail	quest	waist
hoax	raid	
jest	rash	

Long (N=31)		
boredom	lobster	terrorism
bribery	loyalty	torment
certitude	mugging	torture
colonel	peninsula	treason
council	poverty	trial
counsel	prosperity	tributary
estuary	quarrel	twinkle
forgery	smuggling	valley
herring	summit	verdict
jury	tedium	witness
kernel		

Long (N=32)		
account	famine	plumber
animosity	hatred	pollen
ascent	invasion	prejudice
assent	jeopardy	query
bias	landing	sensor
blister	menace	signet
carpenter	mortgage	tailor
controversy	mussel	tornado
custom	peril	whisker
disdain	pigeon	willow
evidence		
excess		

Concrete (N=29)			**Concrete (N=32)**		
cape	kernel	steak	blister	mane	sensor
cod	lamb	stream	bruise	mussel	signet
colonel	lobster	summit	bud	peel	tailor
cove	pane	toe	carpenter	pigeon	tee
ebb	peninsula	tributary	creek	plumber	tool
estuary	prawn	trout	dough	pollen	tornado
flow	ridge	valley	flare	porch	tube
gorge	shrimp	veal	fur	rash	waist
ham	squid	witness	hail	rein	whisker
herring	stake		jug	sage	willow
			landing		
			leek		
Abstract (N=31)			**Abstract (N=28)**		
boredom	jury	spat	account	excess	peril
bribery	loyalty	tedium	animosity	famine	plea
certitude	mugging	terrorism	ascent	hatred	prejudice
council	pain	torment	assent	hoax	query
counsel	poverty	torture	bane	invasion	quest
doubt	proof	tow	bias	jeopardy	raid
forgery	prosperity	treason	controversy	jest	scent
gleam	punch	trial	custom	loan	soul
glee	quarrel	twinkle	disdain	menace	
gloom	smuggling	verdict	evidence	mortgage	
jab					

As we said earlier, we used the General Service List to identify the frequency of the words used in this study and classify them into frequent and infrequent words. Another useful aspect is to check the English words' frequency in the BNC in relation to the Greek equivalents' frequency in the Greek National Corpus. I would like now to comment on the British and the Greek corpus, respectively.

The British National Corpus[35] (BNC) is a 100 million word collection of samples of written and spoken language from a wide range of sources, designed to represent a wide cross-section of current British English, both spoken and written. The corpus comprises 100,106,008 words. The Corpus is designed to represent as wide a range of modern British English as possible. The written part (90%) includes, for example, extracts from

[35] A corpus is a collection of linguistic data, either written texts or a transcription of recorded speech, which can be used as a starting-point of linguistic description or as a means of verifying hypotheses about a language (Crystal, 1991). A corpus can be thought of as a collection of texts gathered according to particular principles for some particular purpose. A corpus allows researchers, teachers and learners to use great amounts of real data in their study of language instead of having to rely on intuitions and made-up examples (Schmitt, 2000:68). For example, frequency (how frequently any particular word occurs in written or spoken language) and collocation (the tendency of two or more words to co-occur in discourse) have been studied almost exclusively through corpus evidence.

newspapers, specialist periodicals and journals for all ages and interests, academic books and popular fiction, published and unpublished letters and memoranda, school and university essays, among many other kinds of text. The spoken part (10%) includes a large amount of unscripted informal conversation, recorded by volunteers selected from different age, region and social classes in a demographically balanced way. In order to find the frequency of the words used in the study we used the BNC Online Service (http://www.natcorp.ox.ac.uk/). We registered for an account to get a free trial of BNC Online Service. We searched the frequencies of the 120 words used in the study (see Table 5.5).

The Hellenic National Corpus (HNC) has been developed by the Institute of Language and Speech Processing (ILSP). The goal of ILSP is to support the growth of Language Technology in Greece and carry out applied research in Speech Processing, Text Processing and Language Learning Technologies. The HNC currently contains more than 34,000,000 words of written texts, so it is smaller than the BNC Corpus (this makes HNC not so reliable for measuring word frequency). Texts in the ILSP Corpus represent Modern Greek language use. They were all published after 1976, most of them having been written after 1990. In order to include different types of language, texts from several media (e.g. books, periodicals or newspapers), belonging to different genres (e.g. non-fiction, advertising) and dealing with various topics (leisure, geography) have been selected. The ILSP Corpus contains samples of written language exclusively. Oral samples have not yet been incorporated (for more information on the ILSP Corpus visit the following website http://hnc.ilsp.gr/en/info.htm). In order to find the frequency of the words used in the study we had access to statistical data concerning the contents of the HNC by clicking on the link 'Statistical data' on the Queries page. Users can have access to certain statistical information including word, lemma and part of speech frequencies. Users can look for specific word or lemma frequencies. The system gives data concerning the 100 and 1,000 most frequent words, lemmas and parts of speech (http://hnc.ilsp.gr/en/statistics.asp).

The following table (Table 5.5) presents the frequencies of the words used in the study both in the BNC and the HNC. We standardized occurrences per 1,000 words. Table 5.5 lists the words in order of frequency starting from the most frequent (the first and last

twenty are in bold). Having a closer look at the table we notice that in the twenty most frequent words both languages have seven in common (*council-simvoulio, doubt-amfivolia, trial-diki, loan-danio, soul-psihi, proof-apodiksi* and *summit-korifi*), while in the twenty less frequent, they have five in common (*leek-praso, tributary-parapotamos, signet-sfragidolithos, mussel-midi* and *blister-fouskala*). Since the common frequency of both languages is relatively small we regard the fact as an insignificant factor in our study[36]. We have to mention here that for the Greek translation of the English words we used the *Oxford Learner's Pocket Dictionary, English-Greek and Greek-English* by D.N.Stavropoulos (1991). It is also important to point out that since a word could have multiple translations in L1 (e.g. *evidence* can be translated in two Greek equivalents *martiria* and *apodiksi*) only one is chosen for the convenience of the study.

Table 5.5: Frequencies of words in rank order from highest to lowest frequency

	English Words	**Frequency in BNC**	Standardized occurrences/ scores per 1,000 words		**Greek Translation of English Words**	**Frequency in HNC**	Standardized occurrences/ scores per 1,000 words
1st	**council**	**31230**	**0,3120**	**1st**	**simvoulio** (council)	**7004**	**0,2060**
2nd	**evidence**	**21166**	**0,2114**	**2nd**	**erotima** (query)	**3605**	**0,1060**
3rd	**account**	**16158**	**0,1614**	**3rd**	**mesi** (waist)	**2976**	**0,0875**
4th	**doubt**	**11907**	**0,1189**	**4th**	**korifi** (summit)	**2163**	**0,0636**
5th	**pain**	**7012**	**0,0700**	**5th**	**kindinos** (peril)	**1827**	**0,0537**
6th	**trial**	**6386**	**0,0638**	**6th**	**apili** (menace)	**1741**	**0,0512**
7th	**flow**	**5192**	**0,0519**	**7th**	**psihi** (soul)	**1719**	**0,0506**
8th	**valley**	**4613**	**0,0461**	**8th**	**diki** (trial)	**1602**	**0,0471**
9th	**loan**	**3812**	**0,0381**	**9th**	**apodiksi** (proof)	**1388**	**0,0408**
10th	**poverty**	**3020**	**0,0302**	**10th**	**anazitisi** (quest)	**1220**	**0,0359**
11th	**soul**	**2909**	**0,0291**	**11th**	**amfivolia** (doubt)	**1016**	**0,0299**
12th	**mortgage**	**2869**	**0,0287**	**12th**	**peristeri** (pigeon)	**1014**	**0,0298**
13th	**excess**	**2808**	**0,0281**	**13th**	**pisti** (loyalty)	**941**	**0,0277**
14th	**proof**	**2636**	**0,0263**	**14th**	**anodos** (ascent)	**819**	**0,0241**
15th	**summit**	**2526**	**0,0252**	**15th**	**danio** (loan)	**815**	**0,0240**
16th	**stream**	**2504**	**0,0250**	**16th**	**tromokratia** (terrorism)	**781**	**0,0230**
17th	**witness**	**2390**	**0,0239**	**17th**	**isvoli** (invasion)	**754**	**0,0222**
18th	**landing**	**2376**	**0,0237**	**18th**	**veveotita** (certitude)	**751**	**0,0221**
19th	**jury**	**2263**	**0,0226**	**19th**	**diamahi** (controversy)	**742**	**0,0218**
20th	**tool**	**2180**	**0,0218**	**20th**	**htipima** (jab)	**735**	**0,0216**
21st	stake	2095	0,0209	**21st**	ergalio (tool)	674	0,0198
22nd	tube	1956	0,0195	**22nd**	martiras (witness)	406	0,0119
23rd	controversy	1926	0,0192	**23rd**	pili (porch)	386	0,0114
24th	invasion	1904	0,0190	**24th**	martiria (evidence)	374	0,0110
25th	colonel	1817	0,0182	**25th**	misos (hatred)	367	0,0108
26th	lamb	1633	0,0163	**26th**	evimeria (prosperity)	340	0,0100

[36] This will be further examined through the Spearman Correlation test conducted and analysed in the next chapter.

27th	loyalty	1602	0,0160	**27th**	lampsi (twinkle)	340	0,0100
28th	punch	1478	0,0148	**28th**	apati (hoax)	323	0,0095
29th	raid	1457	0,0146	**29th**	listia (mugging)	312	0,0092
30th	custom	1447	0,0145	**30th**	pagida (jeopardy)	298	0,0088
31st	ham	1426	0,0142	**31st**	ponos (pain)	295	0,0087
32nd	verdict	1398	0,0140	**32nd**	prodosia (treason)	251	0,0074
33rd	bias	1397	0,0140	**33rd**	enstasi (plea)	249	0,0073
34th	waist	1337	0,0134	**34th**	mirodia (scent)	230	0,0068
35th	prejudice	1299	0,0130	**35th**	etimigoria (verdict)	230	0,0068
36th	ridge	1290	0,0129	**36th**	simvouli (counsel)	227	0,0067
37th	counsel	1284	0,0128	**37th**	logariasmos (account)	221	0,0065
38th	plea	1141	0,0114	**38th**	epidromi (raid)	206	0,0061
39th	prosperity	1112	0,0111	**39th**	lathreborio (smuggling)	180	0,0053
40th	fur	1063	0,0106	**40th**	sigatathesi (assent)	176	0,0052
41st	hatred	1009	0,0101	**41st**	martirio (torment)	175	0,0051
42nd	cape	971	0,0097	**42nd**	grothia (punch)	168	0,0049
43rd	scent	931	0,0093	**43rd**	sidagmatarhis (colonel)	154	0,0045
44th	gloom	913	0,0091	**44th**	perifronisi (disdain)	152	0,0045
45th	torture	851	0,0085	**45th**	dahtilo (toe)	149	0,0044
46th	quest	845	0,0084	**46th**	maragos (carpenter)	143	0,0042
47th	peel	739	0,0074	**47th**	kilada (valley)	140	0,0041
48th	trout	709	0,0071	**48th**	ethimo (custom)	129	0,0038
49th	terrorism	689	0,0069	**49th**	prokatalipsi (prejudice)	119	0,0035
50th	famine	651	0,0065	**50th**	rema (creek)	111	0,0033
51st	toe	648	0,0065	**51st**	tzami (pane)	111	0,0033
52nd	peninsula	615	0,0061	**52nd**	exthrotita (animosity)	94	0,0028
53rd	query	611	0,0061	**53rd**	plastografia (forgery)	87	0,0026
54th	boredom	564	0,0056	**54th**	dorodokia (bribery)	85	0,0025
55th	quarrel	555	0,0055	**55th**	ipothiki (mortgage)	77	0,0023
56th	estuary	524	0,0052	**56th**	moustaki (whisker)	74	0,0022
57th	jug	519	0,0052	**57th**	arni (lamb)	71	0,0021
58th	porch	516	0,0052	**58th**	pliksi (boredom)	66	0,0019
59th	spat	505	0,0050	**59th**	halazi (hail)	59	0,0017
60th	menace	501	0,0050	**60th**	enorki (jury)	50	0,0015
61st	tee	498	0,0050	**61st**	gouna (fur)	49	0,0014
62nd	ascent	497	0,0050	**62nd**	agaliasi (glee)	46	0,0014
63rd	carpenter	469	0,0047	**63rd**	fotovolida (flare)	42	0,0012
64th	steak	466	0,0047	**64th**	faragi (gorge)	42	0,0012
65th	rash	435	0,0043	**65th**	farmaki (bane)	39	0,0011
66th	bud	426	0,0043	**66th**	katifia (gloom)	38	0,0011
67th	assent	410	0,0041	**67th**	himaros (stream)	37	0,0011
68th	tow	410	0,0041	**68th**	akrotirio (cape)	36	0,0011
69th	pigeon	401	0,0040	**69th**	monotonia (tedium)	34	0,0010
70th	tailor	394	0,0039	**70th**	perisia (excess)	33	0,0010
71st	cod	373	0,0037	**71st**	hersonisos (peninsula)	30	0,0009
72nd	herring	352	0,0035	**72nd**	vasanistirio (torture)	26	0,0008
73rd	smuggling	349	0,0035	**73rd**	kefaloskalo (landing)	24	0,0007
74th	gorge	348	0,0035	**74th**	flouda (peel)	24	0,0007
75th	willow	342	0,0034	**75th**	idravlikos (plumber)	23	0,0007
76th	treason	339	0,0034	**76th**	korifogrami (ridge)	22	0,0006
77th	gleam	338	0,0034	**77th**	analabi (gleam)	20	0,0006
78th	jeopardy	337	0,0034	**78th**	giri (pollen)	20	0,0006
79th	pollen	334	0,0033	**79th**	moshari (veal)	20	0,0006
80th	sage	327	0,0033	**80th**	ormos (cove)	19	0,0006

81st	hail	324	0,0032	**81st**	kanata (jug)	18	0,0005
82nd	dough	318	0,0032	**82nd**	astakos (lobster)	18	0,0005
83rd	torment	311	0,0031	**83rd**	raftis (tailor)	18	0,0005
84th	flare	298	0,0030	**84th**	epireasmos (bias)	16	0,0005
85th	peril	285	0,0028	**85th**	limos (famine)	16	0,0005
86th	ebb	272	0,0027	**86th**	kalamari (squid)	14	0,0004
87th	creek	269	0,0027	**87th**	itia (willow)	14	0,0004
88th	cove	261	0,0026	**88th**	brizola (steak)	13	0,0004
89th	rein	257	0,0026	**89th**	anemostrovilos (tornado)	13	0,0004
90th	lobster	245	0,0024	**90th**	paliria (flow)	12	0,0004
91st	disdain	241	0,0024	**91st**	heti (mane)	12	0,0004
92nd	bribery	232	0,0023	**92nd**	palouki (stake)	12	0,0004
93rd	shrimp	225	0,0022	**93rd**	esthitiras (sensor)	11	0,0003
94th	animosity	223	0,0022	**94th**	zimari (dough)	9	0,0003
95th	kernel	218	0,0022	**95th**	eksanthima (rash)	9	0,0003
96th	plumber	214	0,0021	**96th**	halinari (rein)	9	0,0003
97th	sensor	209	0,0021	**97th**	horato (jest)	8	0,0002
98th	forgery	204	0,0020	**98th**	faskomilia (sage)	8	0,0002
99th	prawn	188	0,0019	**99th**	melania (bruise)	7	0,0002
100th	**bruise**	**185**	**0,0018**	**100th**	**boubouki** (bud)	**7**	**0,0002**
101st	**twinkle**	**176**	**0,0018**	**101st**	**ekvoli** (estuary)	**7**	**0,0002**
102nd	**mane**	**164**	**0,0016**	**102nd**	**psiha** (kernel)	**7**	**0,0002**
103rd	**tornado**	**164**	**0,0016**	**103rd**	**filonikia** (quarrel)	**7**	**0,0002**
104th	**glee**	**162**	**0,0016**	**104th**	**ftohia** (poverty)	**6**	**0,0002**
105th	**hoax**	**155**	**0,0015**	**105th**	**bakaliaros** (cod)	**5**	**0,0001**
106th	**squid**	**153**	**0,0015**	**106th**	**midi** (mussel)	**5**	**0,0001**
107th	**veal**	**140**	**0,0014**	**107th**	**rimoulkisi** (tow)	**5**	**0,0001**
108th	**leek**	**132**	**0,0013**	**108th**	**parapotamos** (tributary)	**5**	**0,0001**
109th	**pane**	**126**	**0,0013**	**109th**	**pestrofa** (trout)	**4**	**0,0001**
110th	**tedium**	**122**	**0,0012**	**110th**	**praso** (leek)	**3**	**0,0001**
111th	**jab**	**121**	**0,0012**	**111th**	**garida** (shrimp)	**3**	**0,0001**
112th	**tributary**	**115**	**0,0011**	**112th**	**fouskala** (blister)	**2**	**0,0001**
113th	**signet**	**112**	**0,0011**	**113th**	**aboti** (ebb)	**2**	**0,0001**
114th	**jest**	**104**	**0,0010**	**114th**	**hiromeri** (ham)	**2**	**0,0001**
115th	**whisker**	**92**	**0,0009**	**115th**	**rega** (herring)	**1**	**0,0000**
116th	**mugging**	**86**	**0,0009**	**116th**	**solinari** (tube)	**1**	**0,0000**
117th	**mussel**	**86**	**0,0009**	**117th**	**karavida** (prawn)	**0**	**0,0000**
118th	**blister**	**74**	**0,0007**	**118th**	**sfragidolithos** (signet)	**0**	**0,0000**
119th	**bane**	**68**	**0,0007**	**119th**	**kavgadaki** (spat)	**0**	**0,0000**
120th	**certitude**	**24**	**0,0002**	**120th**	**ipsomataki** (tee)	**0**	**0,0000**

5.7.3 Tasks to complete before beginning the study

We had three tasks to complete before beginning this study:

i) to obtain official permission to undertake research

ii) to distribute Vocabulary Level Tests and Student's Language Background Questionnaires

iii) and to observe how vocabulary is introduced to the class

First, I will talk about the first sample of children EFL learners and then I will comment on the second sample of adult EFL learners.

a) Children EFL learners

We obtained the official permission to undertake research at the private school of languages on 24/02/2004 (see Appendix 2). It is important to make sure that the students who take part in the study belong to the same level of vocabulary knowledge. In order to identify the students' level of vocabulary knowledge we used a Vocabulary Level Test (Nation, 2001). This test is used to make an estimate of a learner's vocabulary size. This means how many high-frequency words the learners already know. The Vocabulary Levels Test was devised by Paul Nation in the early 1980s as a simple instrument for classroom use by teachers in order to help them develop a suitable vocabulary teaching and learning programme for their students. It is a useful tool for diagnostic vocabulary testing for international students (Read, 2000:118). The Vocabulary Levels Test measures knowledge of words at different frequency levels: e.g. 2,000, 3,000, 5,000, or 10,000. It is fairly quick to take. The subjects in our study were tested on a 1000 'Word Level True/False Test'[37] (see Nation, 2001:412) because it applies to beginners and intermediate level. The test consists of 39 sentences. The students are asked to write T if a sentence is true, N if it is not true and X if they do not understand the sentence. We used this vocabulary test to check if all the students would achieve a similar number of correct answers. The vocabulary level tests were administered to the students on 25/02/2004 (see Appendix 3 – raw scores are also included). We compared the means of the test scores for both classes (Class A and Class B). The results of this testing show that both classes achieved a similar number of correct answers (Class A: Mean = 28, 5 SD = 4, 41 and, Class B: Mean = 29 SD = 4, 07).

The Student's Language Background Questionnaire (administered to the students on 25/02/2004) was made to provide further information about the students' age, level, time spent on studying English, etc. (see Appendix 4). The instructions for the 'Word Level True/False Test' and the Student's Language Background Questionnaire were translated into Greek to make sure that all the questions were clear and comprehensible to the students.

---

[37] A description of the making of this type of test can be found in Nation (1993).

Since the words were obtained from an upper-intermediate coursebook it seemed proper to test L1 awareness in the earliest stage of the first study. One possible confounding factor could be that some of the words taught to the children might have been too sophisticated and their meaning difficult to understand (e.g. *forgery*) even in their L1 (Greek). In order to eliminate this factor and ensure that all participants knew the meaning of these words in their L1 (Greek) we provided the children learners with a list[38] of the words translated in Greek. We asked the participants to tick the words they did not know (see Appendix 5). We distributed the list to the children on 27/02/2004. Few of the participants (11 to 12 years old) ticked some of the Greek words (*sigatathesi-assent, etimigoria-verdict, katifia-gloom, sfragidolithos-signet, prokatalipsi-prejudice*). These words were explained to the students though Greek synonyms or definitions during the teaching process. It should be noted that the number of students who had unknown words in their native language was very small (ten) and also the number of unknown Greek words was very small (five). This means that the study was not affected by unknown L1 vocabulary. We ensured that through the teaching process the students would get taught the meaning of the L1 words they did not know.

In order to monitor how vocabulary is introduced to the class we created a Vocabulary Observation List. We observed one class on 25/02/2004 (the teacher was an English native speaker) and one class on 27/02/2004 (the teacher was a non-native speaker). Appendix 6 provides the two vocabulary observation lists. Because we were interested in teachers' opinion on vocabulary presentation and teaching we conducted informal interviews with them. The interviews took place after the classroom observations (the findings from vocabulary observation lists and interviews are reported in the next chapter – section 6.8.1). We have also created a Post-test Questionnaire (in Greek) to provide further information on students' views concerning the two alternative ways of vocabulary presentation (see Appendix 7 – the findings are also reported in section 6.8.1). The Post-test Questionnaire was administered to the students on 21/05/2004.

---

[38] The list did not contain any Greek non-words because first we assumed the students already knew the words in their L1 and second the objective of this study is to test learnability of new L2 vocabulary and not of L1.

b) Adult EFL learners

We obtained the official permission to undertake research on 22/10/2004 (see Appendix 8). Similarly to the children, vocabulary level tests ('Word Level True/False Tests') were administered to the students on 22/10/2004. Appendix 9 presents a sample of the test administered to adult beginners along with their raw scores. We compared the means of the test scores for both classes (Class A and Class B). The results of this testing show that both classes achieved a similar number of correct answers (Class A: Mean = 10, 76 SD = 3, 19 and, Class B: Mean = 9.20 SD = 2, 95). This indicates that all adult students were at the same level of vocabulary knowledge (beginners with a minimal knowledge on English vocabulary).

The Language Background Questionnaire (administered to the students on 22/10/2004) was the same as the one used in the first study apart from questions number 4, 5, 6, and 8 which were excluded from the questionnaire (see Appendix 10). The new questionnaire includes two alternative questions. We wanted to know if the reasons (and needs) for the participants to study English were personal or professional, and if there were any other members of the family who were learning English. We were not able to observe how vocabulary is introduced to the class because the permission for classroom observation was refused (due to the institution's regulations prohibiting the presence of any person other than the instructor). We gave the participants the same L2 word-checklist (described earlier) as given in the first study. The participants ticked a very small number of words on the list (see Appendix 11). The students did not tick any non-words. The remaining words included all the words we taught the students in the first study. For this reason, we decided to teach adult learners the same vocabulary items (120 words) we taught the EFL intermediate students (the L2 check-list was administered to the students on 22/10/2004).

Similarly with the children group, in order to ensure that all adult-participants knew the meaning of the words in their L1 (Greek) we provided them with the same list of the words translated in Greek. We asked the participants to tick the words they did not know (see Appendix 12). The participants did not tick any of the Greek words. This ensured that all participants had knowledge of all the L1 words on the list. The list was

distributed to the participants on 25/10/2004. A Post-test Questionnaire was not administered to the adult participants due to time restrictions.

5.7.4 Timing and stages of the study

This section describes the timing and the stages of the present study. I will first refer to children EFL learners and then to adult EFL learners. A detailed description of the teaching method and the steps followed in each lesson is provided in section 5.7.6.

a) Children EFL learners

Two intermediate classes participated in this study. The subjects in Class A were taught the association between 60 English words and their Greek equivalents with words that were semantically related (topic-related vocabulary, homonyms, synonyms and antonyms) for a period of three weeks. There were two lessons per week. Each vocabulary lesson lasted for forty-five minutes and took place at the end of the normal class that students attended every Monday and Friday.

At the same time, the subjects in Class B were taught the association between 60 English words and their Greek equivalents with words that were not related semantically. The words were presented in a mixed (unrelated) order. At the end of the third week, an immediate (short-term) vocabulary test was administered to both classes. Two weeks later the subjects in both classes were tested on a long-term vocabulary test. For the next three weeks, Class A was taught the association between English words and their Greek equivalents with the words grouped in a mixed (unrelated) order. The vocabulary items were the same used for Class B. In the meantime, Class B was taught the association between English words and their Greek equivalents with the words grouped in a related fashion (semantically related words). The words were the same words used for Class A. At the end of the third week, an immediate (short-term) vocabulary test was administered to both classes. Two weeks later, the subjects in both classes were tested in a long-term vocabulary test. Table 5.6 describes the timing and the schedule of this study in detail. The purpose for reversing the teaching procedure was to see if the order of vocabulary presentation has any effect on learning. The type of

exercises used for teaching the new vocabulary and which group of words was taught each day is discussed later in this chapter.

Table 5.6: Schedule of research with children EFL learners

WEEK	CLASS A		CLASS B	
	Type of words	**Type of exercises**	**Type of words**	**Type of exercises**
1 01/03-07/03	**Monday:** Topic-related 1 **Friday:** Homonyms	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.	**Monday:** Unrelated **Friday:** Unrelated	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.
2 08/03-14/03	**Monday:** Topic-related 2 **Friday:** Synonyms	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.	**Monday:** Unrelated **Friday:** Unrelated	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.
3 15/03-21/03	**Monday:** Topic-related 3 **Friday:** Antonyms	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.	**Monday:** Unrelated **Friday:** Unrelated	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.
4 22/03-28/03	**Monday:** **Short-term TEST**		**Monday:** **Short-term TEST**	
5 29/03-04/04	**Friday:** **Long-term TEST**		**Friday:** **Long-term TEST**	
6 05/04-11/04	**Easter Holidays**		**Easter Holidays**	
7 12/04-18/04	**Easter Holidays**		**Easter Holidays**	

8 19/04-25/04	**Monday:** Unrelated **Friday:** Unrelated	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.	**Monday:** Topic-related 1 **Friday:** Homonyms	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.
9 26/04-02/05	**Monday:** Unrelated **Friday:** Unrelated	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.	**Monday:** Topic-related 2 **Friday:** Synonyms	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.
10 03/05-09/05	**Monday:** Unrelated **Friday:** Unrelated	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.	**Monday:** Topic-related 3 **Friday:** Antonyms	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.
11 10/05-16/05	**Monday:** **Short-term TEST**		**Monday:** **Short-term TEST**	
12 17/05-23/05	**Friday:** **Long-term TEST**		**Friday:** **Long-term TEST**	

b) Adult EFL learners

The schedule of research for this study was the same as in the first one. The study lasted for ten weeks (from October 25th to December 31st). Table 5.7 describes the timing and the schedule of the study in detail.

Table 5.7: Schedule of research with adult EFL learners

WEEK	CLASS A		CLASS B	
	Type of words	**Type of exercises**	**Type of words**	**Type of exercises**
1 25/10-31/10	**Monday:** Topic-related 1 **Friday:** Homonyms	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.	**Monday:** Unrelated **Friday:** Unrelated	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.
2 01/11-07/11	**Monday:** Topic-related 2 **Friday:** Synonyms	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.	**Monday:** Unrelated **Friday:** Unrelated	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.
3 08/11-14/11	**Monday:** Topic-related 3 **Friday:** Antonyms	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.	**Monday:** Unrelated **Friday:** Unrelated	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.
4 15/11-21/11	**Monday:** **Short-term TEST**		**Monday:** **Short-term TEST**	
5 22/11-28/11	**Friday:** **Long-term TEST**		**Friday:** **Long-term TEST**	
6 29/11-05/12	**Monday:** Unrelated **Friday:** Unrelated	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.	**Monday:** Topic-related 1 **Friday:** Homonyms	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.

7 06/12-12/12	**Monday:** Unrelated **Friday:** Unrelated	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.	**Monday:** Topic-related 2 **Friday:** Synonyms	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.
8 13/12-19/12	**Monday:** Unrelated **Friday:** Unrelated	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.	**Monday:** Topic-related 3 **Friday:** Antonyms	Practice with word cards. Matching words with definitions (L1). Blank filling. Answering questions.
9 20/12-26/12	**Monday:** **Short-term TEST**		**Monday:** **Short-term TEST**	
10 27/12-31/12	**Friday:** **Long-term TEST**		**Friday:** **Long-term TEST**	

5.7.5 Presentation of new vocabulary

Before analysing the teaching procedure in detail, it is worth mentioning that during the presentation of new vocabulary, attention was paid to one particular aspect of knowing a word: the form-meaning connection (for a detailed account on aspects of knowing a word see section 1.6). According to Nation (2001:101), strengthening the form-meaning connection involves having to recall a meaning when seeing or hearing a particular word, or having to recall a spoken or written form when wanting to express a meaning. The subjects must be able to recognise a word and link it to its meaning by using L1 translations. This refers to the receptive knowledge of a word which is the main concern of this study. Translation is often criticised as encouraging the idea that there is an exact equivalence between words in the L1 and L2. However, L1 translation has the advantages of being quick, simple and easily understood (Nation, 2001). L1 translations should not be criticized as being a partial and inadequate representation of the word, but should be seen as a useful step in the cumulative process of learning a word (Nation, 2001:82). There is evidence indicating that, particularly in the first meetings with a word, any explanation should not be complicated. Nation (2001:90) points out that it is worth using the learners' L1 if this will provide a clear and brief explanation.

According to Nation (2000:299), for the word form-word meaning aspect of vocabulary learning, direct learning from word cards is an efficient and highly effective technique. A word card has the foreign word on one side and its L1 translation on the other. This allows the learner to be able to retrieve the meaning of the word from memory. Word cards help with learning the written form of the words, learning the concept of the word and making the connection between form and meaning (Nation, 2001:299). Word cards can be used for both receptive and productive learning. Learning from word cards is a way of quickly increasing vocabulary size through focused intentional learning. Several studies of learning from lists or word cards have shown that for many learners learning is faster if the meaning of the word is conveyed by a first language translation (Nation, 2001:66)[39].

We taught new vocabulary to the students by using the three important general processes that may lead to a word being remembered: noticing, retrieval and generation (see section 2.2.3). In addition, the students were asked to do some vocabulary exercises in order to encourage repetition of the new vocabulary in each lesson. Repetition is essential for vocabulary learning because there is so much to know about a word that one meeting with it is not sufficient to gain this information (Nation, 2001:74).

The teaching procedure and the exercises involved students paying attention to both receptive and productive knowledge and use of the words. From the point of view of receptive knowledge and use, the students have to: a) be able to recognise the words when they are heard; b) be familiar with their written form (so that they are recognized when they are met in reading); and c) know that the words signal a particular meaning in the particular context in which it has just occurred. From the point of view of productive knowledge and use, they students have to be able to write the new words with correct spelling and use the words correctly in a sentence. It should be pointed out that the productive aspect of learning is not our main objective in this study. It was minimally used in order to reinforce the receptive aspect of learning. This means that even though exercises referring to productive knowledge (see *Exercise two* in the next section) were used, it was done not in a way to promote productive learning but rather to assist the

[39] For the values of learning from word cards see Nation, 2001:302.

receptive one. The next section on 'Teaching procedure' provides specific examples of how we taught students both receptive and productive knowledge of new vocabulary.

5.7.6 Teaching procedure

The teaching procedure and methods used in this study were influenced by the learning strategies and teaching techniques analysed in Chapter Two. The teaching methodology was the same for both children and adult groups. Both groups were exposed to the same teaching material. Each lesson lasted for forty-five minutes. The teacher (myself as a practitioner[40]) first introduced the students to the new vocabulary and then elaborated, expanded and consolidated these words into classroom activities (exercises). We followed the same teaching procedure for both related and unrelated words. We have to mention here that during each lesson we kept a register for all absent students and we made sure they were given the words taught during their absence.

Step 1 - Duration: ten minutes (Noticing)**:** At first, the students saw a list of ten English words written on the board. The teacher then read aloud the words one by one (in order to familiarize the students with the pronunciation of the new words) and provided their Greek translations (by writing the Greek equivalent of each word on the board i.e. *priest* = *παπάς*). Pronunciation has a very important role in vocabulary presentation. The teacher introduced words making sure that students knew how they are said. This can help students use the words in speech. Thus, every time the teacher wrote up new words on the board she indicated where the stress in the word is (by underlining the stress, e.g. photograph).

Some words, however, are polysemous or homonymous. For example, the word *landing* has two unrelated senses: a) an area in the house and b) the act of bringing an aircraft to the ground. For this study, we provided one of the main meanings of these words. Nation (2001:79) mentions that experimental evidence shows that simultaneous presentation of a word form and its meaning is best for the first encounter and,

---

[40] It should be noted that during this teaching procedure the participants' regular teacher was not present in the classroom.

thereafter, delayed presentation is best because there is then possibility of effort leading to successful recall.

The students wrote the English word on one side of a card and the meaning (using L1 translation) on the other to encourage recall. Small cards (around 5 x 4 cm) were used so that they could be easily carried around. Research shows that learning is generally better if meaning is written in the learners' L1 (Nation, 2001). The students were encouraged to learn words receptively, i.e. to see the L2 word and recall the meaning using L1 translation.

Step 2 - Duration: fifteen minutes (Retrieval)**:** Each of the students went through the set of cards looking at each foreign word and trying to retrieve its meaning. If the student did not remember the Greek equivalent he or she would turn the card over. The students repeated this process for each of the new words. The teacher ensured that the word cards were used repeatedly by practising the word card strategy with the whole group. The purpose of the repetitions was simply to facilitate learning. Tinkham (1993) found that most learners required five to seven repetitions for the learning of a group of six paired associates. Thus, the teacher went through the set of cards with the students at least five to six times.

The students, then, were asked to say (orally) the Greek translation for each new English word. They had to answer oral questions like '*What is a priest?*'. The students had to say the Greek equivalent. The questions help them instantiate and apply the words. Answering oral questions helps learners to use and negotiate new vocabulary items in dialogically symmetrical discourse. According to *Interaction Hypothesis*[41], this seems to create better conditions for vocabulary acquisition.

Step 3 – Duration: twenty minutes (Generation)**:** During the third phase of the teaching process, the students were asked to do two different exercises to encourage repetition of the new vocabulary in each lesson (as mentioned in section 5.7.5). Students remember best when they have actually done something with the words they are

[41] This technique has its roots in *Interaction Hypothesis* which explains how verbal interaction can create the conditions necessary for acquisition to take place (for a review see Ellis, 1995).

learning. Sökmen (1997) points out that if L2 learners meet the word in different contexts by using a variety of activities, a more accurate understanding of the word's meaning and use will be developed. The same format of exercises was used for both class A and class B:

Exercise One: The students had to match the English words with the Greek equivalents.

E.g. Match the English words on the left with the Greek equivalents on the right.

smuggling	δωροδοκία
terrorism	ένορκοι
forgery	τρομοκρατία
mugging	λαθρεμπόριο
trial	ετυμηγορία
proof	ληστεία
jury	πλαστογραφία
verdict	απόδειξη
witness	δίκη
bribery	μάρτυρας

The use of a simple matching activity like this allows the students more involvement than a presentation led by the teacher.

Exercise Two: This is another good way to make the transition between working with the definitions available and recalling what the words mean on their own. We had a word bank with ten vocabulary words and ten sentences (in English) with blanks. The students had to fill in the gaps with the correct word. When the students encountered an unknown word (within the sentences) the teacher provided the Greek equivalent. Nation (2001:309) mentions that relevant studies (e.g. Griffin, 1992, Dempster, 1987) do not show a striking superiority of sentence context over isolated word but, because of the extra information that a sentence context can provide, it is advisable to use such contexts on word cards wherever possible.

E.g. Fill each gap with the most appropriate word from exercise one. Use each word only once.

a) Mary has been a victim of _____ last week.
b) The _____ is the decision that is given by the _____ at the end of a _____.
c) _____ is the act of offering money to someone in order to persuade them to do something for you.
d) _____ is the use of violence, kidnapping and bombing.
e) John was arrested and charged with drug _____.
f) You must have _____ of residence if you want to get into the building.
g) A _____ to an accident or crime is the person who saw it.
h) The document was a _____!

Appendix 13 includes all the exercises given to both classes (Class A and Class B) in each lesson. We produced most of the sentences in *Exercise Two* based on definitions taken from *Collins Cobuild English Dictionary* (1987) and some examples provided on the *Longman Vocabulary Website* (www.ablongman.com/vocabulary). During the last few minutes of the lesson, the teacher made sure that all the students had the correct answers by giving them an answer sheet for both exercises. The students were asked to go through the exercises and report how many correct answers they got. We did not ask students to study the words out of class because this study does not involve autonomous learning (or learners' strategies of learning) and we wanted to exclude it as much as practically possible (though of course students could not be prevented).

5.7.7 Recent trends in vocabulary testing

Testing L2 vocabulary knowledge should take into account all aspects of what is implied by lexical knowledge (see section 1.6.2.2): knowledge of form, meaning, morphology, syntax, collocations and use in discourse. Depending on what exactly one wants to know about L2 lexical knowledge, one has to select the appropriate materials and adequate procedures to arrive at valid and reliable results.

Before analysing the design of the vocabulary test used in this study, it is worth mentioning some recent trends in vocabulary testing. Vocabulary tests are used by language teachers to monitor students' progress in vocabulary learning and to diagnose areas of weakness in their knowledge of L2 words they are learning. According to Read

(2000:7), there are two contrasting perspectives on the role of vocabulary in language assessment. The first point of view represents the idea that it is sensible to write tests that measure whether learners know the meaning and usage of a set of words, taken as independent (semantic) units (e.g. whether students can distinguish *ship* and *sheep*). This view is reflected in the use of conventional vocabulary tests. The second view is that vocabulary should be assessed in the context of a language-use task, where it interacts with other components of language knowledge (e.g. learners are asked to write a letter of complaint to a hotel manager using appropriate vocabulary knowledge). This view favours the communicative format of tests emphasising contextualisation.

Read (2000:9) has proposed a framework of three dimensions to differentiate this dichotomy (see Table 5.8).

Table 5.8: Dimensions of Vocabulary Assessment (adapted from Read, 2000:9)

Discrete A measure of vocabulary knowledge or use as an independent construct	**Embedded** A measure of vocabulary which forms part of the assessment of some other, larger construct
Selective A measure in which specific vocabulary items are the focus of the assessment	**Comprehensive** A measure which takes account of the whole vocabulary content of the input material (reading/listening tasks) or the test-taker's response (writing/speaking tasks)
Context-independent A vocabulary measure in which the test-taker can produce the expected response without referring to any context	**Context-dependent** A vocabulary measure which assesses the test-taker's ability to take account of contextual information in order to produce the expected response

The dichotomy is reflected in the division of *discrete*, *selective*, and *context-independent* tests on the one hand, and *embedded*, *comprehensive*, and *context-dependent* tests on the other hand. In between, however, there are intermediate forms. The (multiple-choice) cloze test and its derivative known as the C-test is an example of an embedded, selective, and context-dependent test. The C-test is created by choosing several short texts and deleting the second half of every second (or x th) word in each text (Read, 2000:111).

Another classification of vocabulary tests reflects the quantitative/qualitative distinction of lexical knowledge: tests estimating vocabulary size (breadth) and quality (depth) of word knowledge (see section 1.6.2.2). An example of a size-related test is the Vocabulary Levels Test devised by Paul Nation in the early 1980s. It involves the task of matching words and definitions (Read, 2000:118-126). In general, size-related tests take the form of checklists, ranging from very simple yes/no-type lists to computerised, frequency-based lists including non-words (Read, 1997:312).

On the other hand, tests of depth of lexical knowledge are classified into two categories: a) tests attempting to analyse the different aspects of lexical knowledge, and b) 'developmental' tests "identifying levels of knowledge that may be interpreted as stages in the acquisition of the word" (Read, 1997:315). According to Read (1997), due to the difficulty of designing tests that accurately assess the complex multidimensional construct of lexical knowledge, existing tests mainly fall into the second category, using some sort of rating scale, such as the Vocabulary Knowledge Scale (VKS) which has the following five steps:

I) I don't remember having seen this word before.

II) I have seen this word before, but I don't know what it means.

III) I have seen this word before, and I think it means ___
(synonym or translation)

IV) I know this word. It means ___. (synonym or translation)

V) I can use this word in a sentence: ___ (Write a sentence.)
(If you do this section, please also do Section IV.)

The VKS elicitation scale (Paribakht and Wesche, 1997:180, in Read, 2000:133)

Despite being "a workable instrument, allowing coverage of a reasonable number of words" (Read, 1997:317) and "sensitive to increases in vocabulary knowledge that result from reading activities" (Read, 2000:135), various aspects of its validity have been questioned. It is not clear, for example, that the levels in the scale correspond to acquisition stages. In addition, multiple meanings of a word are one aspect of lexical knowledge that the VKS in its present form does not account for. Read (1997)

concludes that a rating scale test requires considerable refinement to improve its validity.

As far as measuring written vocabulary production, Laufer and Nation (1995) propose the Lexical Frequency Profile (LFP) which measures the amount of vocabulary from different frequency levels[42] used by learners in their composition writing. The LFP does not show how well particular words are known but it has been shown to be a reliable and valid measure (Laufer and Nation, 1995) which can measure change in language proficiency (Laufer, 1994).

Word association tests have been devised to test (receptive) depth of knowledge. As Read mentions (Read, 2000:180), the standard word-association task involves presenting subjects with a set of stimulus words one by one and asking them to say the first related word that comes into their head. Read has developed an alternative format which presents subjects with a stimulus word and a group of other words. Some of the words are related (syntagmatically i.e. collocates for example *edit-film*, paradigmatically i.e. synonyms such as *edit-revise* or analytically representing one aspect or component of the target word e.g. *team-together*) to the stimulus and some not, the task being to identify the related words (Read, 2000:181). Although Read concludes that the test as a whole functioned well, he mentions that it is not clear to what extent the goal of measuring depth of knowledge has been achieved.

5.7.8 Tests used to test vocabulary knowledge

The test used in this study falls into the category of discrete, selective, and context-independent tests discussed earlier. When we want to measure vocabulary knowledge or use as an independent construct separate from other components of language competence (e.g. listening or speaking), we use a discrete test (Schmitt, 2000:173, Read, 2000). The first step in designing a discrete vocabulary test is to define the purpose of the test (Read, 2000:151). The purpose of the vocabulary test used in this study is to test students on their knowledge of the lexical items (the sixty new L2 words) that they have studied during the course (period of three weeks). For our purposes, we need to find a

[42] The first and second thousand words, the University Word List and remaining words.

vocabulary test which will assess learners' knowledge of individual word meaning, in other words, test the learners' breadth of vocabulary knowledge (see section 1.6.2.1).

An achievement test is the most appropriate vocabulary test to use for this study. This type of test assesses how well the learners have mastered a vocabulary skill that they have been taught (Read, 2000:152), i.e. assesses students' knowledge of the lexical items that they have studied during the course. Nation (2001:373) divides this particular test into a) a short-term achievement test which is used to see whether a recently studied group of words has been learned and b) a long-term achievement which is used to see whether a course (or a specific learning technique) has been successful in teaching particular words in terms of long-term retention. A short-term test and a long-term test for both Classes (Class A and Class B) were constructed. Both tests had the same characteristics (see section 5.7.8.3). It is important, though, to mention that in our case the short-term test is the test conducted two days after the teaching procedure of semantically related vocabulary was completed. The same was done for unrelated vocabulary too. This procedure was chosen in order to illustrate the results of teaching in an actual classroom environment where testing is applied after the completion of a certain number of lessons.

5.7.8.1 The design of the vocabulary test

The vocabulary test used in this study was based on the words taught to the students. According to Read (2000:153), this means that the vocabulary assessment took place within a course of study. The lexical items to be assessed were specified in relation to the learning objectives of the students. The test covered the vocabulary items that the students were supposed to have achieved during the course. A bilingual test format (see section 5.7.8.3) was used because the bilingual list is the predominant kind of note kept on new vocabulary by Greek EFL students (Scholfield and Gitsaki, 1996). This test format employs both the target language (L2) and the learners' own language (Read, 2000:167).

5.7.8.2 Selection of target words

According to Read (2000:159), there is no standard approach to the choice of target words for vocabulary testing because that depends on the objectives and the purpose of the assessment. The target words for this study were all the words that were used in the vocabulary-learning phase (sixty words for Class A and sixty words for Class B for both children and adult groups). The words were presented in a mixed fashion. This means that the words were not in alphabetical order or in the same order they were presented to the students during the course. Nation (2001:345) mentions that a good vocabulary test must have plenty of items (around 30 is probably a minimum for a reliable test).

5.7.8.3 Characteristics of the test

There are three main characteristics of the test:

1) It is a pen and paper test taking 15-20 minutes to complete.

2) It tests receptive knowledge (passive recognition test).

3) It tests vocabulary only.

The way we present the target words in a test is related to the purpose of the assessment. The selected target words were presented in isolation because we wanted to assess the students' ability to supply the meaning when given the target word. Meaning and word (written) form were the main types of word knowledge to be tested. We used a *definition recall test*. This means that the students were given a list of English words and asked to write the Greek equivalent (L1 translation). In this test, translation from L2 to L1 will measure the students' receptive vocabulary knowledge because we are interested in one particular aspect of knowing a word: the form-meaning connection. As mentioned earlier, L1 translations should not be criticized as being a partial and inadequate representation of the word, but it should be seen as a useful step in the cumulative process of learning a word (Nation, 2001). Motivated by that, we judged that testing only receptive vocabulary knowledge by requiring L1 translation would provide a clear picture of students' learning of the new L2 words taught.

Although Read (2000) argues that it is unwise to encourage the learners to believe that any word in L2 has a direct synonymous word or phrase in L1, he also points out that

using L1 translations provides a better means for the test-taker to express their understanding of the target language (Read, 2000:169-170). He also mentions that this technique is effective in providing a foundation for further development of vocabulary knowledge in the L2 (Read, 2000:162).

L1 translations provide a very useful means of testing vocabulary, both receptively and productively, and in recall and recognition items (Nation, 2001:351). This is useful for our study because low-level (in our case beginner and intermediate) learners lack the communicative resources in the target language to be able to express their understanding of the meaning of L2 words through that language (Read, 2000:168). This type of test was easy to mark; it simplified the test-takers' task; and it covered a large sample of words. However, this means that the criterion for 'knowing' a word is a minimal one, and we therefore have to be cautious in interpreting the results of this test. It is based on the measurement of single-meaning senses, and so should be seen as providing only estimates of how many words are *known to a partial extent*.

The students were not informed that they would be tested in two weeks because this would affect the results of the long-term test. As was mentioned before, we constructed a short-term and a long-term test for both classes. During the long-term test, the students were tested on the same words they were tested on the short-term test, but the words were presented in a different order (appendices from 14 to 21 include the short-term and long-term tests for class A and class B for both children and adults).

5.8 Summary

Chapter Five contains a thorough explanation of the complete process surrounding this study. Starting from the initial motivation to test in a real classroom environment the results of previous research (like Tinkham's, 1993, 1997, Waring's, 1997, Schneider, Healy and Bourne's, 1998), an 'action research' with elements of quasi-experiment was conducted. The research design along with the research procedure are fully explained in detail. Factors regarding the selection of Greek EFL students as subjects, timing and stages of the study, vocabulary used, teaching procedure and testing principles are also discussed. The results obtained form vocabulary tests and questionnaires used in the study are analysed and discussed in the next chapter (Chapter Six).

CHAPTER SIX

Data analysis and discussion

6.1 Introduction

This chapter deals with the statistical analysis of the data obtained from the research described in the previous chapter. There is a brief description of the logic of hypothesis testing. Statistical tests used for the analysis are also reported and explained in detail. The rest of the chapter is divided into different sections which describe the statistical procedures accompanied by useful comments provided by data. The results are reported and discussed in detail in the following sections. They are presented in the following order:

1) Related vs unrelated vocabulary test scores
2) Children vs adults' vocabulary test scores
3) Class A performance vs Class B performance in vocabulary test scores
4) Short-term vocabulary test scores vs long-term vocabulary test scores
5) Test scores analysis within semantically related vocabulary
6) Test scores analysis by word properties
7) Analysis of qualitative data

6.2 Logic of Hypothesis Testing

When interpreting an experimental finding a question arises as to whether the finding could have occurred by chance. Hypothesis testing is a statistical procedure for testing whether chance is a plausible explanation of an experimental finding. The hypothesis that an apparent effect is due to chance is called the null hypothesis and is typically the opposite of the researcher's hypothesis which is often called the alternative hypothesis. The null hypothesis states that there is no difference in the distributions of the sample and the population or that the samples belong to the same population (Hinton, 2004:39).

As we said in Chapter Five, the two main hypotheses of this study are:

Hypothesis a: When tested on their knowledge of 60 new L2 words (after the teaching period of three weeks), subjects will achieve a higher test score when tested on related vocabulary compared to their test score on their knowledge of 60 unrelated words.

Hypothesis b: When tested on their knowledge of 60 new L2 words (after the teaching period of three weeks), subjects will achieve a higher test score when tested on unrelated vocabulary compared to their test score on their knowledge of 60 related words.

Likewise, the null hypothesis states that:

Null Hypothesis: When tested on their knowledge of 60 new L2 words (after the teaching period of three weeks), subjects will not achieve a higher test score when tested on related vocabulary compared to their test score on their knowledge of 60 unrelated words (and vice versa). Teaching related or unrelated vocabulary has no positive effect on the subjects' vocabulary test scores (in terms of achieving better scores).

Hypothesis testing is all about comparing a score with a known distribution. We want to know whether a single score comes from a known distribution or from a different population. More usually we compare two samples of subjects to decide whether they come from the same or different populations. Thus, hypothesis testing involves making a decision concerning whether two distributions are the same or different. Assuming that we want to look at the performance of two samples on vocabulary tests. If the difference is small we might be skeptical of a difference in populations but if the difference is large we might decide that the finding indicates a likely difference in populations. The problem we face is how big a difference must be, before we reject the null hypothesis and decide the samples really do come from populations with different distributions. To make this decision we use a decision criterion, the significance level (Hinton, 2004:108). There are three commonly used levels of significance, known as the 5%, 1% and the 0.1% levels respectively. The '5% level' of significance ($p = \leq 0.05$) (read as "p is equal to or less than 0.05") means that there are 5 chances in a 100 that the difference is due to chance. It is normally the lowest level of significance acceptable, and such a result is often called 'significant'. By setting a significant level at $p = 0.05\%$ we are saying that only 5 per cent of the known distribution lies beyond it (Hinton, 2004:97). If the null hypothesis is rejected, then the alternative to the null hypothesis is

accepted. By common convention, if the probability value is below 0.05 then the null hypothesis is rejected.

6.2.1 Types of error in significance testing

Significance testing is concerned to state the degree of confidence that one can have in the test result, of confidence that a mistake is not made when accepting or rejecting the null hypothesis over the alternative hypothesis. Errors in doing this can take one of two forms, known as Type I and Type II errors (Hinton, 2004:96-97):

Type I, often symbolized as α ('alpha') occurs if the null hypothesis is rejected when it is, in fact, true. That means that we are claiming to have found a significant difference between the population distributions where there is, in fact, none. To limit the risk of Type I errors we set our significance level to $\alpha = 0.05$, giving us a 5 in 100 chance, or smaller, of falsely rejecting the null hypothesis.

Type II, (symbolized as β ('beta')) the null hypothesis is accepted when it is, in fact, false. That means that there is, in fact, a significant difference in the population but that we have failed to find it. It means that the data do not provide strong evidence that the null hypothesis is false. Lack of significance does not support the conclusion that the null hypothesis is true.

6.2.2 The power of a statistical test

As mentioned before, a Type II error can only occur if the null hypothesis is false. If the null hypothesis is false, then the probability of a Type II error is called β. The probability of correctly rejecting a false null hypothesis equals $1-\beta$ and is called power (Hinton, 2004).The power of a statistical test refers to its ability to find a difference in distributions when there really is one. According to Hinton (2004:100), the power of a test is related to three factors that we can control: 1) the size of α, 2) the size of the effect we are looking for and, 3) the size of the samples we select. We will examine the three factors separately:

Factor One: It is more important to avoid Type I than Type II error. Hinton (2004) points out that there is always an element of compromise involved in significance

testing. Thus, what we are doing when we select a particular significance level is trying to minimize the risk of both types of errors, α and β.

Factor Two: The effect size (d) is a standardized measure of the difference between the means in terms of standard deviations units. The power of a test and the effect size are increased when the difference between the means of the distributions is large or when their standard deviations are small (Hinton, 2004).

Factor Three: Increasing the sample size increases the power of the test because we reduce the spread of the distribution (Hinton, 2004:104). By increasing the sample size we increase both the power of the test and our confidence in the findings. However, as Hinton (2004:108) mentions, if we have limited resources, time or access to subjects these restrictions may have priorities.

6.3 The use of t – test

When we want to compare two samples we use the t-test (Hinton, 2004:106). The t-test is used for testing differences between two means. In order to use a t-test, the same variable must be measured in different samples, at different times, or in comparison to a known population mean. The t-test allows a comparison of two samples at a time and examines the effect of one independent variable at a time (Hinton, 2004:112). The more common applications of the t-test are testing the difference between independent samples or testing the difference between dependent samples. A t-test for independent samples is useful when the same variable has been measured in two independent samples and the researcher wants to know whether the difference between sample means is statistically significant. "Independent samples" means that the samples have different people in them and that the people in the different samples have not been matched or paired in any way. It is when each subject contributes a score to only one sample (there is no connection between the subjects in the study) (Hinton, 2004:78). A t-test for related samples or a t-test for dependent means is the appropriate test when the same people have been measured. Related samples involve subjects providing scores for both samples (same subjects are used in two conditions) (Hinton, 2004:78).

Hinton, (2004:85) mentions that in order to perform a t-test, we have to make three assumptions:

1) The known population is normally distributed. The requirement for a normal distribution of the population may be relaxed when the sample size is ≥30 (Central Limit Theorem). This is because the sampling distribution of the mean is approximately normal even where the distribution of the population is not normal, provided the sample size is large.

2) The sample is randomly selected from the (unknown) population (so that we can use sample statistics (X, SD) as unbiased estimates of the population).

3) The variance of the unknown population is approximately the same as the variance of the known population (to allow us to use the sample information to estimate population standard deviations).

As we said, the t-test requires that the population distributions are normal but, as Hinton (2004) argues, it is robust against departures from this assumption. When comparing two means, the validity of the t-test also depends on the equality of the two population standard deviations. In many situations it is reasonable to assume this equality.

6.3.1 The Paired-Samples and the Independent-Samples t-test

The Paired-Samples t-test procedure compares the means of two variables for a single group. It computes the differences between values of the two variables for each case and tests whether the average differs from 0. For example, in a study on high blood pressure, all patients are measured at the beginning of the study, given a treatment, and measured again. Thus, each subject has two measures, often called before and after measures (cause and effect). The Independent-Samples[43] t-test procedure compares means for two groups. A low significance value for the t-test (typically less than 0.05) indicates that there is a significant difference between the two variables. If the confidence interval for the mean difference does not contain zero, this also indicates that the difference is significant. If the significance value is high and the confidence interval

[43] Ideally, for this test, the subjects should be randomly assigned to two groups, so that any difference in response is due to the treatment (or lack of treatment) and not to other factors.

for the mean difference contains zero, then it cannot be concluded that there is a significant difference between the means for the two variables.

6.4 The use of ANOVA

The t-test is limited to examining only two samples at a time, and to examining only one IV (Independent Variable) at a time. If we want to examine more than two samples and more than one IV at a time, we need the F-test or else ANOVA. Analysis of variance is used to test the hypothesis that several means are equal. This technique is an extension of the two-sample t test. One basic characteristic of ANOVA is that it compares variability within groups and between groups.

In addition to determining that differences exist among the means, we may want to know which means differ. When the null hypothesis is rejected, the conclusion is that at least one population mean is different from at least one other mean. However, since the ANOVA does not reveal which means are different from which, we have to use specific tests that detect differences among means. There are two types of tests for comparing means: a priori contrasts and post hoc tests (i.e. Tukey HSD test). Contrasts are tests set up before running the experiment and post hoc tests are run after the experiment has been conducted. We can also test for trends across categories.

6.5 T-test analysis and 'Table of Comparisons'

As already mentioned, our main research question is whether teaching semantically related words together, can help EFL learners to memorise and retain more new L2 words as compared to the number of new L2 words EFL learners can memorise and retain when they are taught semantically unrelated words (together). We also wanted to observe how the long-term tests affect the students' performance on the vocabulary tests. Another aspect we want to examine is the performance of Class A in comparison to Class B (in both children and adult groups). The fourth aspect we want to observe is children's performance in relation to adults' performance. The comparisons we are interested in are presented in Table 6.1, 'Table of Comparisons'. There are 39 comparisons in total and they are divided into four groups depending on the separate

variable (condition) we examine each time. The stars indicate the most important comparisons for discussion.

So, we have:

1) Variable One (V1 = Vocabulary): Related Vocabulary vs Unrelated Vocabulary, which includes 14 different comparisons for children, adults and everyone (all together),

2) Variable Two (V2 = Age): Children vs Adults, which includes 6 comparisons,

3) Variable Three (V3 = Class Quality): Class A vs Class B, which includes 8 different comparisons for both children and adults, and

4) Variable Four (V4 = Time): Short-Term Test (SHT) vs Long-Term Test (LT), which includes 11 different comparisons for both children, adults and everyone (all together).

Our aim is to examine how children, adults and all together behave in the four different conditions described above. The t-test is the most appropriate statistical test for the analysis of the following comparisons because it allows a comparison of two samples at a time and examines the effect of one independent variable at a time.

Table 6.1: Table of comparisons

Variable One: Related Vocabulary vs Unrelated Vocabulary
EVERYONE
Short-Term Test (SHT)
***1**. Everyone (All Children + All Adults), Related, SHT vs Everyone (All Children + All Adults), Unrelated, SHT
Long-Term Test (LT)
***2**. Everyone (All Children + All Adults), Related, LT vs Everyone (All Children + All Adults), Unrelated, LT
CHILDREN
Short-Term Test (SHT)
***3**. All Children (class A+B), Related, SHT vs All Children (class A+B), Unrelated, SHT
4. Class A, Related, SHT vs Class A, Unrelated, SHT
5. Class B, Related, SHT vs Class B, Unrelated, SHT
Long-Term Test (LT)
***6**. All Children (class A+B), Related, LT vs All Children (class A+B), Unrelated, LT
7. Class A, Related, LT vs Class A, Unrelated, LT
8. Class B, Related LT vs Class B, Unrelated LT

ADULTS

Short-Term Test (SHT)

***9**. All Adults (class A+B), Related, SHT vs All Adults (class A+B), Unrelated, SHT
10. Class A, Related, SHT vs Class A, Unrelated, SHT
11. Class B, Related, SHT vs Class B, Unrelated, SHT

Long-Term Test (LT)

***12**. All Adults (class A+B), Related, LT vs All Adults (class A+B), Unrelated, LT
13. Class A, Related, LT vs Class A, Unrelated, LT
14. Class B, Related LT vs Class B, Unrelated LT

Variable Two: Children vs Adults

***15**. All Children (class A+B), All Vocabulary (related+unrelated), SHT vs All Adults (class A+B), All Vocabulary (related+unrelated), SHT
***16**. All Children (class A+B), All Vocabulary (related+unrelated), LT vs All Adults (class A+B), All Vocabulary (related+unrelated), LT
17. All Children (class A+B), Related, SHT vs All Adults (class A+B), Related, SHT
18. All Children (class A+B), Related, LT vs All Adults (class A+B), Related, LT
19. All Children (class A+B), Unrelated, SHT vs All Adults (class A+B), Unrelated, SHT
20. All Children (class A+B), Unrelated, LT vs All Adults (class A+B), Unrelated, LT

Variable Three: Class A vs Class B

CHILDREN

***21**. Class A, SHT, Related vs , Class B, SHT, Related
***22**. Class A, LT, Related vs , Class B, LT, Related
***23**. Class A, SHT, Unrelated vs , Class B, SHT, Unrelated
***24**. Class A, LT, Unrelated vs , Class B, LT, Unrelated

ADULTS

***25**. Class A, SHT, Related vs , Class B, SHT, Related
***26**. Class A, LT, Related vs , Class B, LT, Related
***27**. Class A, SHT, Unrelated vs , Class B, SHT, Unrelated
***28**. Class A, LT, Unrelated vs , Class B, LT, Unrelated

Variable Four: Short-Term Test (SHT) vs Long-Term Test (LT)

EVERYONE

***29**. Everyone (All Children + All Adults), All Vocabulary (related+unrelated), SHT vs Everyone (All Children + All Adults), All Vocabulary (related+unrelated), LT

CHILDREN

***30**. All Children (class A+B), All Vocabulary (related+unrelated), SHT vs All Children (class A+B), All Vocabulary (related+unrelated), LT
31. Class A, Related, SHT vs Class A, Related, LT
32. Class A, Unrelated, SHT vs Class A, Unrelated, LT
33. Class B, Related, SHT vs Class B, Related, LT
34. Class B, Unrelated, SHT vs Class B, Unrelated, LT

ADULTS

*35. All Adults (class A+B), All Vocabulary (related+unrelated), SHT vs All Adults (class A+B), All Vocabulary (related+unrelated), LT
36. Class A, Related, SHT vs Class A, Related, LT
37. Class A, Unrelated, SHT vs Class A, Unrelated, LT
38. Class B, Related, SHT vs Class B, Related, LT
39. Class B, Unrelated, SHT vs Class B, Unrelated, LT

6.5.1 Description and summary of the t-test results tables

In this part, I will describe and summarize the t-test results tables for the four variables (conditions) described above in four separate sections.

Section A): Is there a difference in test scores for related vs unrelated vocabulary?

First, I will describe and present the results from the t-test analysis of the performance of everyone (all together) and, children and adults (separately) on related and unrelated vocabulary, both on short (SHT) and long-term (LT) tests. As we mentioned before, the Independent-Samples t-test procedure compares means for two groups of cases. The t-test results of all fourteen comparisons are presented in Appendix 22 (under Variable One section) in detail. Table 6.2 presents the results of the most important pairs for discussion. The output below displays the number of cases N (number of students and number of words), mean value (of students' test scores), standard deviation, standard error and the obtained t value for each comparison. The output also shows the degrees of freedom (df) and probability (2-tailed significance). Sometimes we find that the calculated t has a minus sign. This simply indicates that the mean of Group 1 is smaller than the mean of Group 2.

Table 6.2: Group statistics for related vs unrelated vocabulary

	Groups	N of students	N of words	Mean	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
[1] Everyone-re-SHT vs	**1,00**	**63**	**60**	**21,1746**	**9,48435**	**1,19492**	**-2,501**	**124**	**,014**
Everyone-un-SHT	**2,00**	**63**	**60**	**25,6984**	**10,78176**	**1,35837**			
[2] Everyone-re-LT vs	**1,00**	**63**	**60**	**16,8095**	**7,97346**	**1,00456**	**-3,114**	**124**	**,002**
Everyone-un-LT	**2,00**	**63**	**60**	**21,7460**	**9,73515**	**1,22651**			
[3] All-Children-re-SHT vs	1,00	31	60	18,8065	9,96467	1,78971	-,620	60	,538
All-Children-un-SHT	2,00	31	60	20,4194	10,51277	1,88815			
[6] All-Children-re-LT vs	1,00	31	60	14,0323	7,24101	1,30052	-1,106	60	,273
All-Children-un-LT	2,00	31	60	16,3871	9,39034	1,68655			
[9] All-Adults-re-SHT vs	**1,00**	**32**	**60**	**23,4688**	**8,53072**	**1,50803**	**-3,469**	**62**	**,001**
All-Adults-un-SHT	**2,00**	**32**	**60**	**30,8125**	**8,40675**	**1,48612**			
[12] All-Adults-re-LT vs	**1,00**	**32**	**60**	**19,5000**	**7,82469**	**1,38322**	**-4,032**	**62**	**,000**
All-Adults-un-LT	**2,00**	**32**	**60**	**26,9375**	**6,90459**	**1,22057**			

In comparisons numbered [1], [2], [9] and [12] the p-value is below the '5% level' of significance ($p<0, 05$). This indicates that there is a significant difference between the two variables in each pair. The results allow us to reject the null hypothesis (at the $p = 0.05$ level of significance) and conclude that everyone and adults separately performed significantly better at the 0.05 significance level on the unrelated vocabulary test compared to their performance on the related vocabulary test. It becomes clear from the table above that adults' scores influenced overall (everyone's) performance, since children show no significant difference in test scores between related and unrelated vocabulary. This suggests tentatively that unrelated vocabulary may assist learning of new L2 words more than related vocabulary only at beginners' level (adults). These findings complement the evidence of previous research by Tinkham (1993, 1997), Waring (1997), Schneider, Healy and Bourne (1998) and, Finkbeiner and Nicol (2003) illustrating that presenting L2 students (beginners) with their new vocabulary grouped together in sets of syntactically and semantically similar new words impedes rather than facilitates the learning of the words.

'Interference Theory' supports the results found above claiming that when words are being learned at the same time, but are too 'similar' or share too many common elements, then these words will interfere with each other thus impairing retention of them. Extensive research into interference theory (see Baddeley, 1990) suggests that as similarity increases between targeted information and other information learned either before or after the targeted information the difficulty of learning and remembering the

targeted information also increases (Tinkham, 1993). Similarly the 'distinctiveness hypothesis' (see Hunt and Mitchell, 1982), which relates ease of learning to the distinctiveness (non-similarity) of the information to be learned, also validates the above argument.

It is important to point out that these results apply to beginning level EFL adults and not to intermediate EFL children where there is no significant difference between related vs unrelated vocabulary test scores. According to the high means noticed in adults' performance especially in unrelated vocabulary we made the following assumptions. One probable reason that might influence adults to achieve higher scores is motivation. Motivation has to do with the emotional dimension of learning a L2. It is possible that emotions affect how successful a L2 learner is (Archibald, 1997:526, in O'Grady, Dobrovolsky and Katamba eds.). There are two types of motivation: *instrumental* (for a specific goal, e.g. English language certificate) and *integrative* (for personal reasons, e.g. to learn more about a culture). Both those types simultaneously affected adults' performance since the main reason they joined the English seminars was to acquire a certificate in English in order to use it professionally and for personal interest (information obtained from Language Background Questionnaires - see Appendix 10). Laufer and Hulstijn (2001) point out that all learners and teachers know that motivation promotes success and achievement in L2 learning. This means that students, who have high levels of motivation, will achieve high levels of proficiency.

Another possible reason for adults' higher scores is that adults can master certain aspects of a foreign language even well into adulthood. Scovel (1988)[44] finds no evidence to support the idea that there is a critical period for the acquisition of syntax or lexicon. Adult L2 learners routinely achieve high levels of proficiency in these aspects of a foreign language. Lexical and syntactical competence becomes easier for them in contrast to phonology which becomes very difficult to acquire (Scovel, 1988:123). Even though initially it would seem a paradox having adults (beginners) achieving high scores, the reference above suggests that it could be easier for them to learn new vocabulary.

---

[44] In Bialystok , "Language and Understanding" Brown, Malmkjaer, Pollitt and Williams eds. (1995:123).

Moreover, looking at intermediate subjects' (children's) means of performance we notice quite low scores both in related and unrelated vocabulary. One possible confounding factor for their performance is that the words used in the study were obtained from an upper-intermediate coursebook. This particular source was one level higher than children's intermediate level. Correct though it was to choose this particular source in order to examine their performance in words they did not know, the fact that it may have hindered them is a possibility, due to some words being too advanced for them.

Another reason that might have affected children's low performance is lack of motivation. Keeping in mind motivation as mentioned above (Archibald, 1997:526, in O'Grady, Dobrovolsky and Katamba, eds.) for adults, with children things are different. According to English language knowledge criteria set by private institutions of foreign languages they will be eligible to sit the exams for the Cambridge First Certificate in two years. So, due to time interference, children may not be motivated enough at the present stage.

Additional reasons for children's lower scores can be derived from teachers' interviews (see section 6.8.1). Teachers claim that students do not spend much time studying vocabulary which is caused by lack of motivation and interest. We were also informed that teachers do not spend much time on vocabulary in class due to emphasis on grammar. Another interesting finding from their interviews is that most of the students' time is occupied by attending and studying for Greek state school. This results in lack of concentration and interest in L2 learning.

The impression created by the data analysis reviewed above is that unrelated vocabulary helps beginners (adults), while semantically related does not. Intermediate level students, however, do not seem to be affected by the way vocabulary is presented (related or unrelated). This partly extends the results of previous research which were only limited to beginners (Tinkham, 1993, 1997, Waring, 1997, Schneider, Healy and Bourne, 1998). Additionally, it answers our **first research question** (see section 5.6) and satisfies Waring's suggestion for further research in order to examine performance of intermediate students (Waring, 1997:269), which does not seem to be affected by word presentation (related or unrelated) in a real classroom environment. Regarding

adult beginners though, it solidifies the results of previous research that semantically related vocabulary can impede learning. In contrast, unrelated words are proven to facilitate learning. It is important to mention that these results reinforce the positions stated by the researchers mentioned above since they were extracted from natural language in an EFL classroom through teaching procedure. This is a step forward in trying to remove artificiality from the situation. Additional observations were made about performance of each group (children and adults) separately. There we noticed the high scores of adult students in opposition to the low scores of children learners. Even though possible explanations were given for those observations a clearer picture is demonstrated in the analysis of Variable Two (see below).

Section B): Is there a difference in test scores between children and adults?

In this section, I will describe and present the results from the t-test analysis of the performance of children and adults on short and long-term vocabulary tests (on both related and unrelated vocabulary). Table 6.3 presents the group statistics for variable four: the number of cases N (number of students and number of words), mean value (of students' test scores), standard deviation, standard error and the obtained t value for each comparison. The output also shows the degrees of freedom (df) and probability (2-tailed significance). The t-test results of all comparisons are presented in Appendix 22 (under Variable Two) in detail.

Table 6.3: Group statistics for children vs adults

	Groups	N of students	N of words	Mean	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
[15] All-Children-All-Voc-SHT vs	1,00	62	120	19,6129	10,19057	1,29420	-4,359	124	,000
All-Adults-All-Voc-SHT	2,00	64	120	27,1406	9,18050	1,14756			
[16] All-Children-All-Voc-LT vs	1,00	62	120	15,2097	8,40011	1,06682	-5,408	124	,000
All-Adults-All-Voc-LT	2,00	64	120	23,2188	8,22398	1,02800			
[17] All-Children-re-SHT vs	1,00	31	60	18,8065	9,96467	1,78971	-1,997	61	,050
All Adults-re-SHT	2,00	32	60	23,4688	8,53072	1,50803			
[18] All-Children-re-LT vs	1,00	31	60	14,0323	7,24101	1,30052	-2,876	61	,006
All Adults-re-LT	2,00	32	60	19,5000	7,82469	1,38322			
[19] All-Children-un-SHT vs	1,00	31	60	20,4194	10,51277	1,88815	-4,341	61	,000
All Adults-un-SHT	2,00	32	60	30,8125	8,40675	1,48612			
[20] All-Children-un-LT vs	1,00	31	60	16,3871	9,39034	1,68655	-5,092	61	,000
All Adults-un-LT	2,00	32	60	26,9375	6,90459	1,22057			

All the comparisons are above $p \leq 0.05$, so we conclude that in all cases there is a statistical difference, leading us to the conclusion that adults performed better than children on both short and long-term vocabulary tests (on both related and unrelated vocabulary). Bearing in mind the data analysis of Variable One which presents high mean scores in adults' performance and low mean scores in children's, the present result is expected. The statistical analysis of Table 6.3 clarifies the significant difference between the two groups. As we have already discussed extensively in the analysis of Variable One, adults were highly motivated and more conscious learners for personal and professional reasons. They have also the advantage of being able to master certain aspects of L2 in adulthood (Scovel, 1988). Children had no immediate and clear motivation because of lack of interest. In addition, teachers provided information claiming that vocabulary teaching was overshadowed by grammar presentation and that students did not devote time to studying due to lack of concentration and interest caused by the time they spend on homework for the Greek state schools.

Section C): Is there a difference in test scores between Class A and Class B for children and adults separately?

In this section, I will describe and present the results from the t-test analysis of the performance of Class A and Class B (for both children and adults) on short and long-term vocabulary tests (on both related and unrelated vocabulary). Table 6.4 presents the results from the Independent-Samples t-test procedure. The output below displays the number of cases N (number of students and number of words), mean value (of students' test scores), standard deviation, standard error and the obtained t value for each comparison. The output also shows the degrees of freedom (df) and probability (2-tailed significance). The t-test results of all comparisons are presented in Appendix 22 (under Variable Three section) in detail. In this case, we do not expect a significant difference assuming that class A and class B (in adults and children separately) belong to the same level. This means that in adults' groups both classes are beginners, while in children's both classes are intermediate.

Table 6.4: Group statistics for Class A vs Class B for children and adults

	Groups	N of students	N of words	Mean	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
[21] Children-A-re-SHT vs	1,00	16	60	18,1875	11,11887	2,77972	-,352	29	,727
Children-B-re-SHT	2,00	15	60	19,4667	8,91120	2,30086			
[22] Children-A-re-LT vs	1,00	16	60	13,5000	7,50111	1,87528	-,417	29	,680
Children-B-re-LT	2,00	15	60	14,6000	7,16938	1,85113			
[23] Children-A-un-SHT vs	1,00	16	60	17,7500	12,95376	3,23844	-1,489	29	,147
Children-B-un-SHT	2,00	15	60	23,2667	6,34110	1,63726			
[24] Children-A-un-LT vs	1,00	16	60	15,8750	12,29024	3,07256	-,309	29	,760
Children-B-un-LT	2,00	15	60	16,9333	5,14735	1,32904			
[25] Adults-A-re-SHT vs	1,00	17	60	23,6471	8,47748	2,05609	,124	30	,902
Adults-B-re-SHT	2,00	15	60	23,2667	8,88391	2,29382			
[26] Adults-A-re-LT vs	1,00	17	60	19,8235	8,14122	1,97454	,245	30	,808
Adults-B-re-LT	2,00	15	60	19,1333	7,71702	1,99253			
[27] Adults-A-un-SHT vs	1,00	17	60	30,1765	8,16422	1,98011	-,450	30	,656
Adults-B-un-SHT	2,00	15	60	31,5333	8,90318	2,29879			
[28] Adults-A-un-LT vs	1,00	17	60	27,0000	6,60492	1,60193	,054	30	,958
Adults-B-un-LT	2,00	15	60	26,8667	7,46292	1,92692			

By looking at the table, we notice that in all cases the p-value is above the '5% level' of significance (p<0, 05). This indicates that there is no statistically significant difference between the performance of Class A and Class B (for both children and adults) on short and long-term vocabulary tests (on both related and unrelated vocabulary). Both classes performed in the same way. This strongly validates the homogeneity of the samples and the consistency of the present study.

Section D): Is there a difference in test scores between short-term tests (SHT) and long-term tests (LT)?

In this section, I will describe and present the results from the t-test analysis of the performance of everyone (all together) and, children and adults (separately) on short-term (SHT) and long-term (LT) vocabulary tests. In this case we will use a Paired Samples t-test, because it compares the means of two variables for a single group. All students provided a score for the SHT vocabulary test and then they are tested again on the same vocabulary test in two weeks time (LT test). We want to see if there is a difference in students' performance between SHT and LT tests.

Table 6.5 displays the number of cases N (number of students and number of words), mean value , standard deviation SD, standard error for the pair(s) of variables compared in the paired-samples t-test procedure, the obtain value for t, the degrees of freedom (df) and probability (2-tailed significance). The t-test results of all comparisons are presented in Appendix 22 (under Variable Four section) in detail. Here we present the most important comparisons.

Table 6.5: Group statistics for short-term (SHT) vs long-term tests (LT)

	N of students	N of words	Mean	Std. Deviation	Std. Error Mean	t	df	Sig. (2-tailed)
[29] Everyone (Children + Adults)-All-Vocabulary-SHT	63	120	23,4365	10,36494	,92338	12,877	125	,000
Everyone (Children + Adults)-All-Vocabulary-LT	63	120	19,2778	9,20229	,81981			
[30] All-Children-All-Vocabulary-SHT	31	120	19,6129	10,19057	1,29420	8,214	61	,000
All-Children-All-Vocabulary-LT	31	120	15,2097	8,40011	1,06682			
[35] All-Adults-All-Vocabulary-SHT	32	120	27,1406	9,18050	1,14756	10,632	63	,000
All-Adults-All-Vocabulary-LT	32	120	23,2188	8,22398	1,02800			

From the table above and the results in Appendix 22 we see that there is a significant difference between the means of all pairs (comparisons). In all cases the p-value is below the '5% level' of significance ($p<0.05$). As expected, there is a significant difference between the two variables in each pair. The results reveal that everyone (children and adults), children (class A, class B and all together) and adults (class A, class B and all together) in both related and unrelated vocabulary tests did perform significantly better on short-term tests compared to their performance on long-term tests. This is an effect of memory attrition. As explained in Chapter Five, this study is focused on receptive knowledge of the words. This means that subjects are required to recall the word form providing the L2 translation. It is more likely for a word form and meaning to be recalled in the short term rather than in the long term. As Schmitt (2000:129) mentions, words are not necessarily learned in a linear manner and forgetting is a natural fact of learning. Both learning and forgetting occurs until the word is mastered and 'fixed' in memory. He also points out, that most of the forgetting occurs with words that are only known receptively.

6.6 Analysis of Semantically Related Vocabulary

As we said before (see Chapter Five, section 5.7.2.2), the semantically related nouns (used in the study) were divided into six groups of ten words (three groups of topic-related nouns, one group of homonyms, one group of synonyms and one group of antonyms). The three topic-related groups were based on three different topics; Group 1: crime, Group 2: nature and Group 3: food. We want to check if there is any statistical difference in students' (both children's and adults') performance on semantically related vocabulary (in both short and long-term vocabulary tests). We created the following 'Semantically Related Vocabulary Scores Table' which only describes the raw scores and the means of each group (see Table 6.6). The first column depicts the six groups of semantically related vocabulary used in the study. The second column portrays the number of adults (out of all adults, N=32) who correctly remembered each word in the short-term vocabulary test. The third column presents the number of children (out of all children, N=32) who correctly remembered each word in the short-term vocabulary test while the fourth column provides the total number of both adults and children. The last three columns refer to the long-term vocabulary test results.

Table 6.6: Semantically Related Vocabulary Scores Table

	SHT			**LT**		
	All Adults (N=32)	All Children (N=31)	Everyone (N=63)	All Adults (N=32)	All Children (N=31)	Everyone (N=63)
Homonyms						
pane	6	8	14	3	5	8
pain	22	28	50	23	27	50
steak	7	25	32	7	19	26
stake	6	15	21	4	12	16
toe	16	29	45	12	28	40
tow	13	12	25	12	6	18
colonel	8	8	16	5	5	10
kernel	9	7	16	7	3	10
council	20	13	33	18	11	29
counsel	13	11	24	11	4	15
MEAN	**12**	**15,6**	**27,6**	**10,2**	**12**	**22,2**
Synonyms						
torment	14	6	20	13	4	17
torture	16	6	22	12	4	16
jab	10	2	12	7	1	8
punch	14	6	20	8	5	13
spat	6	0	6	5	0	5
quarrel	16	0	16	14	0	14
gleam	6	3	9	5	1	6

twinkle	11	2	13	10	1	11
boredom	16	2	18	15	2	17
tedium	8	1	9	6	0	6
MEAN	**11,7**	**2,8**	**14,5**	**9,5**	**1,8**	**11,3**
Antonyms						
ebb	14	18	32	10	13	23
flow	10	14	24	8	10	18
gloom	15	10	25	11	8	19
glee	12	8	20	11	5	16
certitude	10	15	25	8	13	21
doubt	10	9	19	10	7	17
loyalty	12	8	20	12	8	20
treason	16	6	22	13	3	16
poverty	10	12	22	10	9	19
prosperity	14	10	24	13	5	18
MEAN	**12,3**	**11**	**23,3**	**10,6**	**8,1**	**18,7**
Topic-related						
Group 1						
smuggling	15	8	23	11	2	13
terrorism	15	10	25	15	7	22
forgery	15	10	25	14	6	20
mugging	4	18	22	4	9	13
trial	25	18	43	23	15	38
proof	22	7	29	20	4	24
jury	18	11	29	17	11	28
verdict	11	5	16	6	3	9
witness	27	27	54	26	24	50
bribery	14	6	20	11	4	15
MEAN	**16,6**	**12**	**28,6**	**14,7**	**8,5**	**23,2**
Group 2						
cape	10	6	16	5	6	11
peninsula	15	14	29	11	12	23
cove	8	24	32	4	20	24
tributary	2	4	6	1	2	3
valley	14	17	31	12	16	28
gorge	14	4	18	11	4	15
stream	11	7	18	8	6	14
estuary	8	7	15	5	3	8
ridge	6	7	13	1	3	4
summit	10	4	14	6	3	9
MEAN	**9,8**	**9,4**	**19,2**	**6,4**	**7,5**	**13,9**
Group 3						
lamb	20	25	45	20	26	46
herring	8	3	11	6	1	7
veal	10	4	14	9	2	11
ham	14	6	20	11	6	17
cod	8	4	12	8	4	12
trout	13	6	19	10	4	14
prawn	19	1	20	17	0	17
shrimp	8	4	12	6	2	8
squid	5	4	9	5	2	7
lobster	12	7	19	10	5	15
MEAN	**11,7**	**6,4**	**18,1**	**10,2**	**5,2**	**15,4**

We observe from the above table high mean scores in Homonyms (Everyone: 27, 6 in SHT and 22, 2 in LT) and Topic-related 1 (Everyone: 28, 6 in SHT and 23, 2 in LT). This may indicate that homonyms and topic-related 1 (crime) could facilitate learning of new L2 vocabulary. On the other hand, we notice low mean scores in Synonyms (Everyone: 14, 5 in SHT and 11, 3 in LT). This observation may suggest that synonymy hinders L2 vocabulary learning. For better evaluation of the above scores we created the Facility Value Table of Semantically Related Vocabulary in descending order from 'easy' to 'difficult' (see Table 6.7).

Table 6.7 provides the facility values for all semantically related words in descending order. These are the scores that all children and adults obtained in both short and long-term tests. By looking at the words that more than 55% of the students have recalled during the testing phase (both SHT and LT test) we notice that two words (*toe, pain*) belong to the 'Homonyms' category, two words (*witness, trial*) in the 'Topic 1' (crime) group and one word (*lamb*) in 'Topic 3' category (food). On the contrary, if we look at the words that few subjects (less than 10% of the students) have recalled, we observe that three words (*spat, gleam, tedium*) belong to the 'Synonyms' category and two words (*tributary, ridge*) are included in the 'Topic 2' (nature) group. We notice that homonyms and topic 1 may be more easily recalled, while synonyms and topic 2 may hinder students' performance. In addition, it is important to mention that the results were not influenced by the time (date) of teaching. For example, *witness* and *trial* were taught on the first day of the teaching phase. However, this is the descriptive side of the analysis of these results. We only have indications whether a certain group would actually impede or facilitate students' performance. We are interested to see if there is a statistically significant difference between students' performance and the six different groups of semantically related vocabulary. This requires a statistical analysis of Table 6.6 (see below).

Table 6.7: Facility Value Table of Semantically Related Vocabulary in descending order from 'easy' to 'difficult'

	WORDS	Everyone (N=63) SHT	Standardized Scores	WORDS	Everyone (N=63) LT	Standardized Scores
1	witness	54	0,86	pain	50	0,79
2	pain	50	0,79	witness	50	0,79
3	toe	45	0,71	lamb	46	0,73
4	lamb	45	0,71	toe	40	0,63
5	trial	43	0,68	trial	38	0,60
6	council	33	0,52	council	29	0,46
7	steak	32	0,51	jury	28	0,44
8	ebb	32	0,51	valley	28	0,44
9	cove	32	0,51	steak	26	0,41
10	valley	31	0,49	proof	24	0,38
11	proof	29	0,46	cove	24	0,38
12	peninsula	29	0,46	ebb	23	0,37
13	jury	29	0,46	peninsula	23	0,37
14	tow	25	0,40	terrorism	22	0,35
15	terrorism	25	0,40	certitude	21	0,33
16	gloom	25	0,40	loyalty	20	0,32
17	forgery	25	0,40	forgery	20	0,32
18	certitude	25	0,40	gloom	19	0,30
19	prosperity	24	0,38	poverty	19	0,30
20	flow	24	0,38	tow	18	0,29
21	counsel	24	0,38	flow	18	0,29
22	smuggling	23	0,37	prosperity	18	0,29
23	treason	22	0,35	torment	17	0,27
24	torture	22	0,35	boredom	17	0,27
25	poverty	22	0,35	doubt	17	0,27
26	mugging	22	0,35	ham	17	0,27
27	stake	21	0,33	prawn	17	0,27
28	torment	20	0,32	stake	16	0,25
29	punch	20	0,32	torture	16	0,25
30	prawn	20	0,32	glee	16	0,25
31	loyalty	20	0,32	treason	16	0,25
32	ham	20	0,32	counsel	15	0,24
33	glee	20	0,32	bribery	15	0,24
34	bribery	20	0,32	gorge	15	0,24
35	trout	19	0,30	lobster	15	0,24
36	lobster	19	0,30	quarrel	14	0,22
37	doubt	19	0,30	stream	14	0,22
38	stream	18	0,29	trout	14	0,22
39	gorge	18	0,29	punch	13	0,21
40	boredom	18	0,29	smuggling	13	0,21
41	verdict	16	0,25	mugging	13	0,21

42	quarrel	16	0,25	cod	12	0,19
43	kernel	16	0,25	twinkle	11	0,17
44	colonel	16	0,25	cape	11	0,17
45	cape	16	0,25	veal	11	0,17
46	estuary	15	0,24	colonel	10	0,16
47	veal	14	0,22	kernel	10	0,16
48	summit	14	0,22	verdict	9	0,14
49	pane	14	0,22	summit	9	0,14
50	twinkle	13	0,21	pane	8	0,13
51	ridge	13	0,21	jab	8	0,13
52	shrimp	12	0,19	estuary	8	0,13
53	jab	12	0,19	shrimp	8	0,13
54	cod	12	0,19	herring	7	0,11
55	herring	11	0,17	squid	7	0,11
56	tedium	9	0,14	**gleam**	**6**	**0,10**
57	squid	9	0,14	**tedium**	**6**	**0,10**
58	gleam	9	0,14	**spat**	**5**	**0,08**
59	**tributary**	**6**	**0,10**	**ridge**	**4**	**0,06**
60	**spat**	**6**	**0,10**	**tributary**	**3**	**0,05**

The statistical analysis of the 'Semantically Related Vocabulary Scores Table' (Table 6.6) will be done in two stages. In the first stage, we will examine if there is a statistical difference between students' performance (adults, children) within semantically related groups (horizontally as we look at the Table) for both the short and the long-term vocabulary test. In other words, we want to examine students' (children's and adults') performance in homonyms, synonyms, antonyms and topic-related vocabulary separately. In the second stage, we will check students' performance (adults, children and everyone) across semantically groups (vertically as we look at the Table). In this case, we compare homonyms, synonyms, antonyms and topic-related vocabulary with each other in relation to students' scores in each group. We want to see how students performed in one group (e.g. synonyms) in relation to the rest of the groups of semantically related vocabulary. Since we want to examine more than two samples and one independent variable at a time, we need the F-test or else ANOVA (analysis of variance). More specifically we will use the One Factor (One way) ANOVA.

Stage 1: Statistical analysis within semantically related groups (horizontally)

In Stage 1, we will compare students' (children's and adults') performance in homonyms, synonyms, antonyms and topic-related vocabulary separately in order to see

if there is any significant difference between students' scores. The ANOVA results of all comparisons are presented in Appendix 23 (under Stage 1 section) in detail. Table 6.8 presents the statistically significant pairwise comparisons of adults' and children's scores (both in SHT and LT tests) within the semantically related group of synonyms. The following table displays descriptive statistics for each group and the pairwise comparisons of the group means. N indicates the size of each group and the number of students. Mean shows the average values. The standard deviation indicates the amount of variability of the scores in each group. The Mean difference lists the differences between the sample means. The output also shows standard error for each group (and pairwise comparisons). Sig lists the probability that the population mean difference is zero.

Table 6.8: ANOVA for comparisons of adults' and children's scores in synonyms

		N of words	N of students	Mean	SD	Std. Error	Mean Difference	Sig.
Synonyms	Adults (Short term)	10	32	11,70	4,057	1,283		
	Children (Short term)	10	31	2,80	2,394	,757		
	Adults (Short term) vs Children (Short term)	10	63			1,418	8,90(*)	,000
	Adults (Long term)	10	32	9,50	3,808	1,204		
	Children (Long term)	10	31	1,80	1,874	,593		
	Adults (Long term) vs Children (Long term)	10	63			1,418	7,70(*)	,000

In the table above, the mean difference results marked with an asterisk (*) indicate a statistical difference due to the significance being lower than 0, 05. The results point out that adults performed better in synonyms than children in both short and long-term tests. One possible reason for the result presented above is that adults' L1 lexicon is more advanced and well developed having stronger semantic links between the words. The L1 lexical entry in the mental lexicon is considered to contain semantic, syntactic, morphological and formal (phonological and orthographic) specifications about a lexical item (Jiang, 2000:48). One important aspect of L1 lexical representation is that these different types of information are highly integrated within each entry and automatically become accessible (Jiang, 2000:49). This integration requires a high and extensive exposure to the language. L1 adult learners are able to extract semantic, syntactic and morphological information while becoming acquainted with the form of

the word, because there is a highly contextualised input. As Jiang (2000) points out, at the initial stage, the use of the L2 word activates the links between L2 words and their L1 translations. In receptive use of the language, the recognition of a L2 word activates its L1 translation equivalent, concerning semantic, syntactic, and morphological information which then becomes available and assists comprehension. Keeping in mind that synonymy belongs to the semantic aspect of L1 lexical representation, this may have allowed adult L2 learners to prevail in synonyms over children.

Stage 2: Statistical analysis across semantically related groups (vertically)

As we mentioned earlier, in the second stage we will we compare homonyms, synonyms, antonyms and topic-related vocabulary with each other in relation to students' (adults', children's and everyone's) scores in each group. A full and detailed description of the results appears in Appendix 23 (under Stage 2 section). Table 6.9 presents the statistically significant pairwise comparisons of students' performance (adults', children's and all together) across semantically related groups for both short and long-term vocabulary tests. N indicates the size of each group and the number of students. Mean shows the average values. The standard deviation indicates the amount of variability of the scores in each group. The Mean difference lists the differences between the sample means. The output also shows standard error for each group (and pairwise comparisons). Sig lists the probability that the population mean difference is zero.

Table 6.9: ANOVA for comparisons of homonyms, synonyms and topic-related in relation to students scores

		N of words	N of students	Mean	SD	Std. Error	Mean Difference	Sig.
Short term / Children	Homonyms	10	31	15,60	8,514	2,692		
	Synonyms	10	31	2,80	2,394	,757		
	Topic-related (3)	10	31	6,40	6,753	2,135		
	Topic-related (1)	10	31	12,00	6,928	2,191		
	Homonyms vs Synonyms	10	31			2,899	12,80(*)	,001
	Homonyms vs Topic-related (3)	10	31			2,899	9,20(*)	,033
	Topic-related (1) vs Synonyms	10	31			2,899	9,20(*)	,033

Long term / Adults	Topic-related (1)	10	32	14,70	7,056	2,231		
	Topic-related (2)	10	32	6,40	4,006	1,267		
	Topic-related (1) vs Topic-related (2)	10	32			2,536	8,30(*)	,025
Long term / Children	Homonyms	10	31	12,00	9,487	3,000		
	Synonyms	10	31	1,80	1,874	,593		
	Homonyms vs Synonyms	10	31			2,927	10,20(*)	,013
Short term / Everyone	Topic-related (1)	10	63	28,60	11,481	3,631		
	Synonyms	10	63	14,50	5,503	1,740		
	Topic-related (1) vs Synonyms	10	63			4,407	14,10(*)	,031

In the table above, the mean difference results marked with an asterisk (*) indicate a statistical difference due to the significance being lower than 0.05. So we observe distinction in semantically related vocabulary comparisons within a group and specific test (long or short term). Children in both short and long-term tests have performed better in homonyms than synonyms. They also achieved higher scores in topic-related 1 (crime) than synonyms (in short-term test). Furthermore, we notice that they performed better in homonyms than topic-related 3 (food). It seems possible that synonyms hinder children while homonyms and topic-related 1 (crime) help them. The superiority of homonymy vs synonymy leads us to assume that phonology may assist children more than multiplicity of word meaning. Homonyms are word forms pronounced or spelt in the same way but having unrelated senses far apart from each other and not obviously related to each other in any way. One claim is that the basis of the operations of the L2 lexicon is phonological rather than semantic (Laufer, 1989:17 – referring to Fromkin, 1971; Hatch, 1983; Soudek, 1982). Phonology appears to play a much more prominent organizing role in the L2 mental lexicon than semanticity[45]. Motivated by this, it is possible to say that children perform better receptively with phonology rather than meaning connection. L2 learners at a low[46] level of proficiency register vocabulary more by phonological similarities than by semantic relatedness (Henning, 1973, in Singleton, 1999:154).

---

[45] A detailed discussion on this matter can be found in Chapter One, section 1.5.3 and 1.7

[46] Compared to a highly advanced EFL student, an intermediate young learner could be considered as a 'low' level student.

Also the fact that topic 1 (crime) seems to facilitate children (in contrast to synonyms) maybe has to do with L2 speakers' exposure to violence and crime. The 'Input Hypothesis' states that language is picked up or acquired when learners receive input from 'messages' which contain language a little above their existing understanding and from which they can infer meaning (Hedge, 2000:10 refering to Krashen, 1985). Children are subjected to crime-related vocabulary through English speaking movies with Greek subtitles (which often appear on Greek television or cinema). This type of exposure may also be the reason that adults and everyone (both adults and children) performed better in topic-related 1 (crime) than topic-related 2 (nature) and synonyms respectively.

6.7 Test scores analysis by word properties

As we mentioned before, we made sure that all nouns used in the study fulfil the criteria for selecting words discussed in Chapter Five, section 5.7.2.2. The total number of words used in the study is 120. The list includes 31 frequent and 89 infrequent words, 57 short and 63 long words, as well as 61 concrete and 59 abstract. It would be interesting to examine if word frequency, word length and word concreteness/abstractness affect all students' performance on vocabulary tests. The main question that this section will try to answer is **'What properties of words influence test scores as a measure of ease of recall in a word translation test when these words are presented in related or unrelated manner?'**. Table 6.10 demonstrates the semantically related and unrelated vocabulary in descending order. These are the scores that all children and adults (everyone) provided in both short and long-term tests. We created this table in order to gain a preliminary insight into the influence of word properties on test scores. By looking at the words that more than 60% of the students have recalled during the testing phase (both SHT and LT tests), we notice that in semantically related vocabulary both in SHT and LT tests we have five words in common (demonstrated in bold). We observe that four out of five are frequent (*witness, pain, toe, trial*). On the contrary, if we look at the words that few subjects (less than 10% of the students) have recalled in the LT test we observe that four out of five are infrequent (*gleam, tedium, ridge, tributary*). So, the first indication we get is that word frequency may facilitate receptive learning. Respectively, in unrelated vocabulary we see that the top ten words are all frequent. This supports the conclusion made about the

role of frequency and allows us to extend it to unrelated vocabulary as well. Thus word frequency seems to be important, regardless of relatedness or unrelatedness. In addition, after examining word length, concreteness and abstractness, we do not get any sufficient results in order to support the importance of these properties. These factors along with frequency will be further examined in the following sub-sections in order to determine their impact on students' scores.

Table 6.10: Facility Value Table of All Vocabulary in descending order from 'easy' to 'difficult'

	LIST OF WORDS	Everyone N=63			LIST OF WORDS	Everyone N=63	
	RELATED	SHT	Standardized Scores		RELATED	LT	Standardized Scores
1st	witness	54	0,86		pain	50	0,79
2nd	pain	50	0,79		witness	50	0,79
3rd	lamb	45	0,71		lamb	46	0,73
4th	toe	45	0,71		toe	40	0,63
5th	trial	43	0,68		trial	38	0,60
6th	council	33	0,52		council	29	0,46
7th	cove	32	0,51		jury	28	0,44
8th	ebb	32	0,51		valley	28	0,44
9th	steak	32	0,51		steak	26	0,41
10th	valley	31	0,49		cove	24	0,38
11th	jury	29	0,46		proof	24	0,38
12th	peninsula	29	0,46		ebb	23	0,37
13th	proof	29	0,46		peninsula	23	0,37
14th	certitude	25	0,40		terrorism	22	0,35
15th	forgery	25	0,40		certitude	21	0,33
16th	gloom	25	0,40		forgery	20	0,32
17th	terrorism	25	0,40		loyalty	20	0,32
18th	tow	25	0,40		gloom	19	0,30
19th	counsel	24	0,38		poverty	19	0,30
20th	flow	24	0,38		flow	18	0,29
21st	prosperity	24	0,38		prosperity	18	0,29
22nd	smuggling	23	0,37		tow	18	0,29
23rd	mugging	22	0,35		boredom	17	0,27
24th	poverty	22	0,35		doubt	17	0,27
25th	torture	22	0,35		ham	17	0,27
26th	treason	22	0,35		prawn	17	0,27
27th	stake	21	0,33		torment	17	0,27
28th	bribery	20	0,32		glee	16	0,25
29th	glee	20	0,32		stake	16	0,25
30th	ham	20	0,32		torture	16	0,25
31st	loyalty	20	0,32		treason	16	0,25
32nd	prawn	20	0,32		bribery	15	0,24
33rd	punch	20	0,32		counsel	15	0,24
34th	torment	20	0,32		gorge	15	0,24

35th	doubt	19	0,30		lobster	15	0,24
36th	lobster	19	0,30		quarrel	14	0,22
37th	trout	19	0,30		stream	14	0,22
38th	boredom	18	0,29		trout	14	0,22
39th	gorge	18	0,29		mugging	13	0,21
40th	stream	18	0,29		punch	13	0,21
41st	cape	16	0,25		smuggling	13	0,21
42nd	colonel	16	0,25		cod	12	0,19
43rd	kernel	16	0,25		cape	11	0,17
44th	quarrel	16	0,25		twinkle	11	0,17
45th	verdict	16	0,25		veal	11	0,17
46th	estuary	15	0,24		colonel	10	0,16
47th	pane	14	0,22		kernel	10	0,16
48th	summit	14	0,22		summit	9	0,14
49th	veal	14	0,22		verdict	9	0,14
50th	ridge	13	0,21		estuary	8	0,13
51st	twinkle	13	0,21		jab	8	0,13
52nd	cod	12	0,19		pane	8	0,13
53rd	jab	12	0,19		shrimp	8	0,13
54th	shrimp	12	0,19		herring	7	0,11
55th	herring	11	0,17		squid	7	0,11
56th	gleam	9	0,14		**gleam**	**6**	**0,10**
57th	squid	9	0,14		**tedium**	**6**	**0,10**
58th	tedium	9	0,14		**spat**	**5**	**0,08**
59th	**spat**	**6**	**0,10**		**ridge**	**4**	**0,06**
60th	**tributary**	**6**	**0,10**		**tributary**	**3**	**0,05**
	UNRELATED				**UNRELATED**		
1st	**custom**	**49**	**0,78**		**account**	**44**	**0,70**
2nd	**account**	**48**	**0,76**		**custom**	**43**	**0,68**
3rd	**hatred**	**46**	**0,73**		**tailor**	**42**	**0,67**
4th	**pigeon**	**45**	**0,71**		**pigeon**	**41**	**0,65**
5th	**tailor**	**45**	**0,71**		**hatred**	**39**	**0,62**
6th	**tornado**	**42**	**0,67**		fur	36	0,57
7th	**fur**	**41**	**0,65**		waist	35	0,56
8th	**waist**	**41**	**0,65**		soul	34	0,54
9th	**soul**	**38**	**0,60**		tornado	33	0,52
10th	**tool**	**38**	**0,60**		bud	32	0,51
11th	bruise	37	0,59		tool	32	0,51
12th	animosity	36	0,57		willow	30	0,48
13th	menace	36	0,57		animosity	29	0,46
14th	mussel	35	0,56		query	29	0,46
15th	evidence	34	0,54		bruise	28	0,44
16th	bud	33	0,52		carpenter	28	0,44
17th	query	33	0,52		evidence	28	0,44
18th	willow	32	0,51		mussel	28	0,44
19th	carpenter	31	0,49		sensor	27	0,43
20th	invasion	31	0,49		menace	26	0,41
21st	porch	31	0,49		tube	26	0,41
22nd	sensor	31	0,49		invasion	25	0,40
23rd	jug	28	0,44		porch	24	0,38
24th	quest	28	0,44		hoax	23	0,37
25th	tube	28	0,44		jug	23	0,37
26th	creek	27	0,43		loan	23	0,37

27th	blister	26	0,41	signet	23	0,37
28th	hoax	26	0,41	creek	22	0,35
29th	jest	26	0,41	jest	22	0,35
30th	mortgage	25	0,40	mortgage	22	0,35
31st	signet	25	0,40	quest	21	0,33
32nd	hail	24	0,38	scent	21	0,33
33rd	famine	23	0,37	plea	20	0,32
34th	loan	23	0,37	tee	20	0,32
35th	scent	23	0,37	excess	19	0,30
36th	plea	22	0,35	leek	19	0,30
37th	prejudice	22	0,35	controversy	18	0,29
38th	tee	22	0,35	prejudice	18	0,29
39th	excess	20	0,32	rein	18	0,29
40th	leek	20	0,32	hail	17	0,27
41st	raid	20	0,32	famine	16	0,25
42nd	jeopardy	19	0,30	pollen	16	0,25
43rd	bias	18	0,29	blister	15	0,24
44th	controversy	18	0,29	peel	15	0,24
45th	mane	18	0,29	sage	15	0,24
46th	pollen	17	0,27	jeopardy	14	0,22
47th	ascent	16	0,25	raid	14	0,22
48th	flare	16	0,25	ascent	13	0,21
49th	peel	16	0,25	bias	13	0,21
50th	rash	16	0,25	mane	12	0,19
51st	rein	16	0,25	rash	12	0,19
52nd	sage	16	0,25	assent	11	0,17
53rd	assent	14	0,22	dough	11	0,17
54th	dough	13	0,21	peril	11	0,17
55th	plumber	13	0,21	plumber	11	0,17
56th	landing	12	0,19	flare	10	0,16
57th	bane	11	0,17	disdain	9	0,14
58th	peril	11	0,17	landing	9	0,14
59th	whisker	9	0,14	whisker	9	0,14
60th	disdain	7	0,11	bane	8	0,13

All the suggestions extracted at first glance from Table 6.10 will be examined using ANOVA in order to look for any significantly statistical difference between test scores and word properties. Since we want to examine more than two samples and more than one independent variable at a time, we need the F-test or else ANOVA (analysis of variance). More specifically, we will use the Two Factor (Two way) ANOVA. This part is divided into three sub-sections: sub-section 1 (6.7.1) refers to subjects' performance in relation to word frequency, sub-section 2 (6.7.2) refers to subjects' performance in relation to word length and sub-section 3 (6.7.3) presents the statistical results of subjects' performance in relation to word concreteness/abstractness (Appendix 24 presents the data for all sub-sections in detail). In all sub-sections, the findings concerning subjects' performance will be presented separately and at the end of each sub-section there will be an overall discussion.

6.7.1 Word frequency

Table 6.11 presents the statistically significant results of adults' performance in relation to word frequency. Table 6.12 presents the statistically significant results of children's performance in relation to word frequency. Table 6.13 presents the statistically significant results of everyone's performance (children and adults together) in relation to word frequency. A full and detailed description of all the results concerning word frequency appears in Appendix 24 (under Word Frequency section). Tables 6.11, 6.12 and 6.13 display descriptive statistics for each group and the pairwise comparisons of the group means. N indicates the size of each group. Mean shows the average values. The standard deviation indicates the amount of variability of the scores in each group. The Mean difference lists the differences between the sample means. The output also shows standard error for each group. Sig lists the probability that the population mean difference is zero.

Table 6.11: ANOVA for comparisons of adults' scores and word frequency

Adults	N of words	N of students	Mean	Std. Deviation	Std. Error	Mean Difference	Sig.
Frequent words / Unrelated / Short term	15	32	18,13	5,290	1,366	6,861(*)	,001
Infrequent words / Related / Short term	44	32	11,27	4,223	,637		
Frequent words / Unrelated / Long term	15	32	16,60	5,040	1,301	7,532(*)	,000
Infrequent words / Related / Long term	44	32	9,07	4,332	,653		
Infrequent words / Unrelated / Short term	45	32	15,11	6,289	,938	3,838(*)	,017
Infrequent words / Related / Short term	44	32	11,27	4,223	,637		
Infrequent words / Unrelated / Long term	45	32	13,09	5,325	,794	4,021(*)	,010
Infrequent words / Related / Long term	44	32	9,07	4,332	,653		

The table above demonstrates that:

1. adults performed better in frequent and unrelated vocabulary than infrequent and related words (in both SHT and LT test)
2. adults performed better in infrequent and unrelated vocabulary than infrequent and related words (in both SHT and LT test)

Table 6.12: ANOVA for comparisons of children's scores and word frequency

Children	N of words	N of students	Mean	Std. Deviation	Std. Error	Mean Difference	Sig.
Frequent words / Unrelated / Short term	15	31	16,47	8,167	2,109	8,035(*)	,001
Infrequent words / Related / Short term	44	31	8,43	6,128	,924		
Frequent words / Unrelated / Short term	15	31	16,47	8,167	2,109	8,289(*)	,001
Infrequent words / Unrelated / Short term	45	31	8,18	6,054	,902		
Frequent words / Unrelated / Long term	15	31	14,20	7,504	1,938	8,382(*)	,001
Infrequent words / Related / Long term	44	31	5,82	5,550	,837		
Frequent words / Unrelated / Long term	15	31	14,20	7,504	1,938	7,956(*)	,002
Infrequent words / Unrelated / Long term	45	31	6,24	5,175	,771		

The table above demonstrates that:

3. children performed better in frequent and unrelated vocabulary than infrequent and related (in both SHT and LT test)
4. children performed better in frequent and unrelated vocabulary than infrequent and unrelated (in both SHT and LT test)

Table 6.13: ANOVA for comparisons of everyone's (children and adults together) scores and word frequency

Everyone	N of words	N of students	Mean	Std. Deviation	Std. Error	Mean Difference	Sig.
Frequent words / Related / Long term	16	63	24,50	13,609	3,402	9,614(*)	,010
Infrequent words / Related / Long term	44	63	14,89	7,701	1,161		
Frequent words / Unrelated / Short term	15	63	34,60	12,052	3,112	14,895(*)	,000
Infrequent words / Related / Short term	44	63	19,70	7,648	1,153		
Frequent words / Unrelated / Short term	15	63	34,60	12,052	3,112	11,311(*)	,001
Infrequent words / Unrelated / Short term	45	63	23,29	8,604	1,283		
Frequent words / Unrelated / Long term	15	63	30,80	10,805	2,790	15,914(*)	,000
Infrequent words / Related / Long term	44	63	14,89	7,701	1,161		
Frequent words / Unrelated / Long term	15	63	30,80	10,805	2,790	11,467(*)	,001
Infrequent words / Unrelated / Long term	45	63	19,33	7,087	1,056		

The table above demonstrates that:

5. everyone performed better in frequent and related vocabulary than infrequent and related in LT test
6. everyone performed better in frequent and unrelated vocabulary than infrequent and related (in both SHT and LT test)
7. everyone performed better in frequent and unrelated vocabulary than infrequent and unrelated (in both SHT and LT test)

At first sight, in 6 out of 7 results above it is stated that students' performance in frequent words is better than infrequent. However, it is important to mention that in some cases we cannot examine frequency separately, since there are frequent and unrelated words compared with infrequent and related (1, 3 and 6). In these cases, words simultaneously have two factors being compared: frequent and unrelated vs infrequent and related. This means that it cannot be clear if frequency affects test scores more than unrelatedness. Keep in mind that as shown in section 6.5.1 unrelated vocabulary is more helpful than related. So, it is not certain if the reason for these results is frequency, or if it simply has to do with the fact that unrelated vocabulary

appears to be more effective than related. In other words, we cannot be sure of the degree to which frequency or unrelatedness affect better performance, as we have seen (see Table 6.2) that unrelated vocabulary may be more helpful.

In result number 3 we have a similar comparison to the ones mentioned above but, taking into consideration that children were not affected by unrelated and related vocabulary (Table 6.2), it would be rather safe to assume that frequency plays the biggest part in performance. In connection to the above, a closer look at results 4, 5, and 7 is acquired. More specifically:

- in results numbers 4 and 7 we have frequent and unrelated compared with infrequent and also unrelated
- in result number 5 we have frequent and related compared with infrequent and also related

In those cases, the effect of frequency appears definite. For example, in result 5, where frequency and relatedness are more efficient than non-frequency and relatedness, it is obvious that since relatedness is common, frequency is the only positive effect on performance. As a result, the influence of frequency may be definite in three cases (4, 5, and 7), while in case number 3 it is a possibility. So, it seems possible that word frequency may make a difference and is worth more attention and research in order to determine its actual impact. For this reason, we performed a Spearman Rank-order Correlation Coefficient. This is a rank-order correlation coefficient which measures association at the ordinal level. It is based on the ranking of the data rather than the actual values. The values of the correlation coefficient (R) range from -1 to 1. The sign of the correlation coefficient indicates the direction of the relationship (positive or negative). The absolute value of the correlation coefficient indicates the strength, with larger absolute values indicating stronger relationships. This means that the closer the value is to 1 the stronger the correlation.

For the requirements of this tests we used Table 6.10 (Facility Value Table of All Vocabulary in descending order from 'easy' to 'difficult') in order to determine the Facility Value Rank Order. This column describes which word comes first in scores, which second, etc. (see Table 6.14 for SHT test scores and Table 6.15 for LT test scores). Since frequency was measured by the BNC, we used the number of appearances in the corpus in rank order. This means that the word with the highest

number of appearances comes first, that with the following number of appearances comes second, etc. (see column Rank Order in BNC in Table 6.14 for SHT test scores and Table 6.15 for LT test scores). The same was done for the Greek translation equivalent from its number of appearances in the HNC (see column Rank Order in HNC in Table 6.14 for SHT test scores and Table 6.15 for LT test scores). We used the Hellenic corpus to see if there is a correlation between test scores and frequency in L1. We also examine the correlation between the two corpora.

Table 6.14: Rank order table for SHT test scores

	LIST OF WORDS	All together N=63				
	ALL	SHT	Standardized scores	Facility Value Rank Order	Rank Order in BNC	Rank Order of Greek equivalent in HNC
1	witness	54	0,86	1	17	22
2	pain	50	0,79	2	5	31
3	custom	49	0,78	3	30	48
4	account	48	0,76	4	3	37
5	hatred	46	0,73	5	41	25
6	lamb	45	0,71	6	26	57
7	pigeon	45	0,71	7	69	12
8	tailor	45	0,71	8	70	83
9	toe	45	0,71	9	51	45
10	trial	43	0,68	10	6	8
11	tornado	42	0,67	11	103	89
12	fur	41	0,65	12	40	61
13	waist	41	0,65	13	34	3
14	soul	38	0,60	14	11	7
15	tool	38	0,60	15	20	21
16	bruise	37	0,59	16	100	99
17	animosity	36	0,57	17	94	52
18	menace	36	0,57	18	60	6
19	mussel	35	0,56	19	117	106
20	evidence	34	0,54	20	2	24
21	bud	33	0,52	21	66	100
22	council	33	0,52	22	1	1
23	query	33	0,52	23	53	2
24	cove	32	0,51	24	88	80
25	ebb	32	0,51	25	86	113
26	steak	32	0,51	26	64	88
27	willow	32	0,51	27	75	87
28	carpenter	31	0,49	28	63	46
29	invasion	31	0,49	29	24	17
30	porch	31	0,49	30	58	23
31	sensor	31	0,49	31	97	93
32	valley	31	0,49	32	8	47
33	jury	29	0,46	33	19	60
34	peninsula	29	0,46	34	52	71
35	proof	29	0,46	35	14	9
36	jug	28	0,44	36	57	81
37	quest	28	0,44	37	46	10
38	tube	28	0,44	38	22	116
39	creek	27	0,43	39	87	50
40	blister	26	0,41	40	118	112
41	hoax	26	0,41	41	105	28
42	jest	26	0,41	42	114	97

43	certitude	25	0,40	43	120	18
44	forgery	25	0,40	44	98	53
45	gloom	25	0,40	45	44	66
46	mortgage	25	0,40	46	12	55
47	signet	25	0,40	47	113	118
48	terrorism	25	0,40	48	49	16
49	tow	25	0,40	49	68	107
50	counsel	24	0,38	50	37	36
51	flow	24	0,38	51	7	90
52	hail	24	0,38	52	81	59
53	prosperity	24	0,38	53	39	26
54	famine	23	0,37	54	50	85
55	loan	23	0,37	55	9	15
56	scent	23	0,37	56	43	34
57	smuggling	23	0,37	57	73	39
58	mugging	22	0,35	58	116	29
59	plea	22	0,35	59	38	33
60	poverty	22	0,35	60	10	104
61	prejudice	22	0,35	61	35	49
62	tee	22	0,35	62	61	120
63	torture	22	0,35	63	45	72
64	treason	22	0,35	64	76	32
65	stake	21	0,33	65	21	92
66	bribery	20	0,32	66	92	54
67	excess	20	0,32	67	13	70
68	glee	20	0,32	68	104	62
69	ham	20	0,32	69	31	114
70	leek	20	0,32	70	108	110
71	loyalty	20	0,32	71	27	13
72	prawn	20	0,32	72	99	117
73	punch	20	0,32	73	28	42
74	raid	20	0,32	74	29	38
75	torment	20	0,32	75	83	41
76	doubt	19	0,30	76	4	11
77	jeopardy	19	0,30	77	78	30
78	lobster	19	0,30	78	90	82
79	trout	19	0,30	79	48	109
80	bias	18	0,29	80	33	84
81	boredom	18	0,29	81	54	58
82	controversy	18	0,29	82	23	19
83	gorge	18	0,29	83	74	64
84	mane	18	0,29	84	102	91
85	stream	18	0,29	85	16	67
86	pollen	17	0,27	86	79	78
87	ascent	16	0,25	87	62	14
88	cape	16	0,25	88	42	68
89	colonel	16	0,25	89	25	43
90	flare	16	0,25	90	84	63
91	kernel	16	0,25	91	95	102
92	peel	16	0,25	92	47	74
93	quarrel	16	0,25	93	55	103
94	rash	16	0,25	94	65	95
95	rein	16	0,25	95	89	96
96	sage	16	0,25	96	80	98
97	verdict	16	0,25	97	32	35
98	estuary	15	0,24	98	56	101
99	assent	14	0,22	99	67	40
100	pane	14	0,22	100	109	51
101	summit	14	0,22	101	15	4
102	veal	14	0,22	102	107	79
103	dough	13	0,21	103	82	94
104	plumber	13	0,21	104	96	75
105	ridge	13	0,21	105	36	76
106	twinkle	13	0,21	106	101	27
107	cod	12	0,19	107	71	105
108	jab	12	0,19	108	111	20
109	landing	12	0,19	109	18	73
110	shrimp	12	0,19	110	93	111

111	bane	11	0,17	111	119	65
112	herring	11	0,17	112	72	115
113	peril	11	0,17	113	85	5
114	gleam	9	0,14	114	77	77
115	squid	9	0,14	115	106	86
116	tedium	9	0,14	116	110	69
117	whisker	9	0,14	117	115	56
118	disdain	7	0,11	118	91	44
119	spat	6	0,10	119	59	119
120	tributary	6	0,10	120	112	108

Table 6.15: Rank order table for LT test scores

	LIST OF WORDS	All together N=63				
	All	LT	Standardized scores	Facility Value Rank Order	Rank Order in BNC	Rank Order of Greek equivalent in HNC
1	pain	50	0,79	1	5	31
2	witness	50	0,79	2	17	22
3	lamb	46	0,73	3	26	57
4	account	44	0,70	4	3	37
5	custom	43	0,68	5	30	48
6	tailor	42	0,67	6	70	83
7	pigeon	41	0,65	7	69	12
8	toe	40	0,63	8	51	45
9	hatred	39	0,62	9	41	25
10	trial	38	0,60	10	6	8
11	fur	36	0,57	11	40	61
12	waist	35	0,56	12	34	3
13	soul	34	0,54	13	11	7
14	tornado	33	0,52	14	103	89
15	bud	32	0,51	15	66	100
16	tool	32	0,51	16	20	21
17	willow	30	0,48	17	75	87
18	animosity	29	0,46	18	94	52
19	council	29	0,46	19	1	1
20	query	29	0,46	20	53	2
21	bruise	28	0,44	21	100	99
22	carpenter	28	0,44	22	63	46
23	evidence	28	0,44	23	2	24
24	jury	28	0,44	24	19	60
25	mussel	28	0,44	25	117	106
26	valley	28	0,44	26	8	47
27	sensor	27	0,43	27	97	93
28	menace	26	0,41	28	60	6
29	steak	26	0,41	29	64	88
30	tube	26	0,41	30	22	116
31	invasion	25	0,40	31	24	17
32	cove	24	0,38	32	88	80
33	porch	24	0,38	33	58	23
34	proof	24	0,38	34	14	9
35	ebb	23	0,37	35	86	113
36	hoax	23	0,37	36	105	28
37	jug	23	0,37	37	57	81
38	loan	23	0,37	38	9	15
39	peninsula	23	0,37	39	52	71
40	signet	23	0,37	40	113	118
41	creek	22	0,35	41	87	50
42	jest	22	0,35	42	114	97
43	mortgage	22	0,35	43	12	55
44	terrorism	22	0,35	44	49	16

45	certitude	21	0,33	45	120	18
46	quest	21	0,33	46	46	10
47	scent	21	0,33	47	43	34
48	forgery	20	0,32	48	98	53
49	loyalty	20	0,32	49	27	13
50	plea	20	0,32	50	38	33
51	tee	20	0,32	51	61	120
52	excess	19	0,30	52	13	70
53	gloom	19	0,30	53	44	66
54	leek	19	0,30	54	108	110
55	poverty	19	0,30	55	10	104
56	controversy	18	0,29	56	23	19
57	flow	18	0,29	57	7	90
58	prejudice	18	0,29	58	35	49
59	prosperity	18	0,29	59	39	26
60	rein	18	0,29	60	89	96
61	tow	18	0,29	61	68	107
62	boredom	17	0,27	62	54	58
63	doubt	17	0,27	63	4	11
64	hail	17	0,27	64	81	59
65	ham	17	0,27	65	31	114
66	prawn	17	0,27	66	99	117
67	torment	17	0,27	67	83	41
68	famine	16	0,25	68	50	85
69	glee	16	0,25	69	104	62
70	pollen	16	0,25	70	79	78
71	stake	16	0,25	71	21	92
72	torture	16	0,25	72	45	72
73	treason	16	0,25	73	76	32
74	blister	15	0,24	74	118	112
75	bribery	15	0,24	75	92	54
76	counsel	15	0,24	76	37	36
77	gorge	15	0,24	77	74	64
78	lobster	15	0,24	78	90	82
79	peel	15	0,24	79	47	74
80	sage	15	0,24	80	80	98
81	jeopardy	14	0,22	81	78	30
82	quarrel	14	0,22	82	55	103
83	raid	14	0,22	83	29	38
84	stream	14	0,22	84	16	67
85	trout	14	0,22	85	48	109
86	ascent	13	0,21	86	62	14
87	bias	13	0,21	87	33	84
88	mugging	13	0,21	88	116	29
89	punch	13	0,21	89	28	42
90	smuggling	13	0,21	90	73	39
91	cod	12	0,19	91	71	105
92	mane	12	0,19	92	102	91
93	rash	12	0,19	93	65	95
94	assent	11	0,17	94	67	40
95	cape	11	0,17	95	42	68
96	dough	11	0,17	96	82	94
97	peril	11	0,17	97	85	5
98	plumber	11	0,17	98	96	75
99	twinkle	11	0,17	99	101	27
100	veal	11	0,17	100	107	79
101	colonel	10	0,16	101	25	43
102	flare	10	0,16	102	84	63
103	kernel	10	0,16	103	95	102
104	disdain	9	0,14	104	15	4
105	landing	9	0,14	105	18	73
106	summit	9	0,14	106	91	44
107	verdict	9	0,14	107	32	35
108	whisker	9	0,14	108	115	56
109	bane	8	0,13	109	119	65
110	estuary	8	0,13	110	56	101
111	jab	8	0,13	111	111	20
112	pane	8	0,13	112	109	51

113	shrimp	8	0,13	113	93	111
114	herring	7	0,11	114	72	115
115	squid	7	0,11	115	106	86
116	gleam	6	0,10	116	77	77
117	tedium	6	0,10	117	110	69
118	spat	5	0,08	118	59	119
119	ridge	4	0,06	119	36	76
120	tributary	3	0,05	120	112	108

The results from Spearman Correlation Coefficient are presented in Table 6.16[47].

Table 6.16: Spearman Correlation for test scores (SHT and LT) and frequencies in corpora

			Facility Value Rank Order of SHT test scores	Rank Order in BNC of SHT test scores	Rank Order of Greek equivalent in HNC of SHT test scores
Spearman's	Facility Value Rank Order of SHT test scores	Correlation Coefficient	1,000	**,308(*)**	**,249(*)**
		Sig. (2-tailed)	.	,001	,007
		N	120	120	116
	Rank Order in BNC of SHT test scores	Correlation Coefficient	**,308(*)**	1,000	**,349(*)**
		Sig. (2-tailed)	,001	.	,000
		N	120	120	116
	Rank Order of Greek equivalent in HNC of SHT test scores	Correlation Coefficient	**,249(*)**	**,349(*)**	1,000
		Sig. (2-tailed)	,007	,000	.
		N	116	116	116
			Facility Value Rank Order of LT test scores	Rank Order in BNC of LT test scores	Rank Order of Greek equivalent in HNC of SHT test scores
	Facility Value Rank Order of LT test scores	Correlation Coefficient	1,000	**,338(*)**	**,257(*)**
		Sig. (2-tailed)	.	,000	,005
		N	120	120	116
	Rank Order in BNC of LT test scores	Correlation Coefficient	**,338(*)**	1,000	**,349(*)**
		Sig. (2-tailed)	,000	.	,000
		N	120	120	116
	Rank Order of Greek equivalent in HNC of SHT test scores	Correlation Coefficient	**,257(*)**	**,349(*)**	1,000
		Sig. (2-tailed)	,005	,000	.
		N	116	116	116

The correlations marked with an asterisk are significant at the 0.01 level (2-tailed). The correlation coefficients on the main diagonal are always 1.0, because each variable has a perfect positive linear relationship with itself. Correlations above the main diagonal are a mirror image of those below. The Spearman's correlation coefficient for:

[47] We used the SPSS 12 to perform the Spearman Correlation test.

- test scores and BNC is 0,308 in SHT and 0,338 in LT test
- test scores and HNC is 0,249 in SHT and 0,257 in LT test
- rank order of appearances in BNC and HNC is 0,349

Since the values above are relatively far from 1, this indicates that the compared variables are weakly correlated. The significance of each correlation coefficient is also displayed in Table 6.16. The significance level (in all cases) is very small (less than 0.05). This means that the correlation is significant and the variables are linearly related. Overall, the results indicate that there is a weak significant positive correlation between test scores and word frequency in language.

According to Nation, the positive role of word frequency in L2 vocabulary retention is expected. Although a language makes use of a large number of words, not all of these words are equally useful. One measure of usefulness is word frequency. This means how often the word occurs in normal use of the language (Nation and Waring, in Schmitt and McCarthy eds., 1997:8). The classic list of high-frequency words is Michael West's (1953) *A General Service List of English Words* which contains approximately 2,000 word families (as mentioned in section 2.3). These words are very important because they account for at least 85% of the words on any page of any book (Nation, 1990). For this reason, Nation suggests that teachers and students should pay attention and spend considerable time with them.

The frequency of a word in language, however important for learning as shown above, is also a usage factor dependent on the type of language input that the learner receives. For example, the frequency of a word's occurrence may be very different in general language (as recorded e.g. in BNC) as compared to the language used in a classroom (Laufer, 1997:141). It is also the case that frequency of occurrence in language input is strongly related to the probability of knowing[48] a word for both L1 and L2 (Vermeer, 2001).

[48] See Chapter One, section 1.6.2.2.

6.7.2 Word length

Table 6.17 presents the only statistically significant results which occurred only in adults' performance in relation to word length. A full and detailed description of all the results concerning word length appears in Appendix 24 (under Word Length section). Table 6.17 displays descriptive statistics for each group and the pairwise comparisons of the group means.

Table 6.17: ANOVA for comparisons of adults' scores and word length

Adults	N of words	N of students	Mean	Std. Deviation	Std. Error	Mean Difference	Sig.
Long words / Unrelated / Short term	32	32	16,59	6,628	1,172	5,042(*)	,010
Short words / Related / Short term	29	32	11,55	4,903	,911		
Long words / Unrelated / Long term	32	32	14,97	5,986	1,058	5,624(*)	,002
Short words / Related / Long term	29	32	9,34	5,246	,974		

Reading the table above, we notice that adults had better results with long and unrelated vocabulary than short and related in both the SHT and LT test. In addition, by looking at Appendix 24 we observe that there is no statistical difference, indicating that word length does not affect the performance of children or students as a whole. Even though at first glance it appears that long words benefit adults' performance in contrast to short ones, unrelatedness vs relatedness should not be neglected. This means that we cannot say with any certainty whether length is more important than unrelated presentation of L2 vocabulary.

Presuming that word length is more powerful than unrelatedness in this case, it is worth discussing why long words seem to assist adult learners more than short. Laufer (1997:145) states that one common misperception is to assume that shorter words are easier because they are more frequent in the language (as explained in section 5.7.2.1). For the purpose of this study long words consist of two or more syllables. In Greek language the majority of nouns contain two or more syllables. This fact means that adults are more exposed to and become more familiar with long words than short. So, since they are more accustomed to learning long words than short in their L1 it is

possible that they may find L2 long words easier than short ones. As Swan (in Schmitt and McCarthy eds. 1997:156) points out, the mother tongue has a considerable influence on the way a L2 is learned and used; this means that L2 characteristics that are similar to the learner's native language will be simple for learners, and those elements that are different will probably be more difficult. It should be noted though that if the L1 was another language in which only short words are frequent, the results above might be different.

6.7.3 Word concreteness/abstractness

Table 6.18 presents the statistically significant results which occurred only in adults' performance in relation to word concreteness/abstractness. A full and detailed description of the results for all groups concerning word concreteness/abstractness appears in Appendix 24 (under Word Concreteness/Abstractness section). Table 6.18 displays descriptive statistics for each group and the pairwise comparisons of the group means.

Table 6.18: ANOVA for comparisons of adults' scores and word concreteness/abstractness

Adults	N of words	N of students	Mean	Std. Deviation	Std. Error	Mean Difference	Sig.
Abstract words / Unrelated / Short term	28	32	18,07	6,230	1,177	7,106(*)	,000
Concrete words / Related / Short term	29	32	10,97	5,186	,963		
Abstract words / Unrelated / Short term	28	32	18,07	6,230	1,177	4,426(*)	,030
Abstract words / Related / Short term	31	32	13,65	4,737	,851		
Abstract words / Unrelated / Long term	28	32	16,18	5,437	1,027	7,627(*)	,000
Concrete words / Related / Long term	29	32	8,55	5,402	1,003		
Abstract words / Unrelated / Long term	28	32	16,18	5,437	1,027	4,308(*)	,039
Abstract words / Related / Long term	31	32	11,87	4,842	,870		

The following conclusions can be drawn from the table:
1) adults performed better in abstract and unrelated vocabulary than concrete and related (in both SHT and LT tests) and
2) adults performed better in abstract and unrelated vocabulary than abstract and related (in both SHT and LT tests).
Furthermore, Appendix 24 points out that word concreteness or abstractness does not affect children's or all subjects' performance.

The second result simply indicates the point made on Variable One (see section 6.5.1) that unrelated vocabulary leads to better performance than related vocabulary, as far as adult students are concerned. This happens because the factor of abstractness is the same for both pairs. The first result, however, implies that abstract words are more important in adults' performance than concrete. Unrelatedness and relatedness are accompanying factors and we cannot determine their degree of influence.

It might be worth examining the possibility of abstract words being more helpful in receptive learning than concrete. Nelson & Shreider[49] (1992), found that subjects perform better on concrete words than on abstract words because a) concrete words are associated with a smaller number of other words than abstract words, making them easier to recall, b) concrete words have larger and densely connected sets, making them easier to recall than abstract words, and c) the 'Imaginability Hypothesis' assumes that subjects are more likely to generate images for concrete than for abstract words. This additional visual storage gives them their edge over abstract words in recognition, recall and lexical decision tasks. However, the first result contradicts all of the above and proves them invalid. One possible reason could be that unrelatedness facilitated adults' performance regardless of the factor of abstractness. This finding verifies the result of Variable One (see section 6.5.1) and strengthens the second result presented in this part which shows that abstractness is unconnected with performance.

[49] Cited in Coady and Huckin (1997:213).

6.7.4 Factorial ANOVA of word properties

In the previous sub-sections we examined the influence of word properties in test scores in relation to word grouping for teaching presentation (related or unrelated). In order to examine the influence of word properties separately (apart from relatedness or unrelatedness) in test scores, we will perform a Factorial ANOVA using GLM Univariate[50]. The General Linear Model (GLM) Univariate procedure provides regression analysis and analysis of variance for one dependent variable (scores, in our case) by one or more factors (e.g. word frequency, in our case). The factor variables divide the population into groups (e.g. frequent and infrequent, in our case). Using this General Linear Model procedure, we can test null hypotheses about the effects of other variables on the means of various groupings of a single dependent variable. We can investigate interactions between factors as well as the effects of individual factors, some of which may be random.

For the requirements of this test we used Table 5.3 (Chapter Five) which presents the words in alphabetical order adding the scores of each word (obtained from Table 6.10). The scores were our dependent variable which were compared with the properties of words. This was done separately for SHT and LT test scores. Table 6.19 presents the results of Factorial ANOVA. The column labelled Source lists the effects in the model. The mean square[51] of each effect is presented in the second column. The F statistic and its significance value are displayed in the next columns. The F statistic is calculated by dividing the mean square by the mean square error. Effects with a small significance value (smaller than 0.05) are significant.

---

[50] I am indebted to Dr Manolis Gavezos for advising me on running a factorial ANOVA.
[51] The mean square of each effect is calculated by dividing the sum of squares by its degrees of freedom.

Table 6.19: Factorial ANOVA for test scores and word properties

Dependent Variable: Test Scores (SHT and LT)

Source	Mean Square	F	Sig.
Word Frequency in SHT	2205,004	22,586	**,000**
Word Concretness in SHT	176,375	1,807	,182
Word Length in SHT	181,891	1,863	,175
Word Frequency in SHT * Word Concretness in SHT	91,143	,934	,336
Word Frequency in SHT * Word Length in SHT	141,290	1,447	,232
Word Concretness in SHT * Word Length in SHT	,173	,002	,966
Word Frequency in SHT * Word Concretness in SHT * Word Length in SHT	18,508	,190	,664
Error	97,627		
Word Frequency in LT	2628,447	30,614	**,000**
Word Concretness in LT	155,967	1,817	,180
Word Length in LT	128,421	1,496	,224
Word Frequency in LT * Word Concretness in LT	95,300	1,110	,294
Word Frequency in LT * Word Length in LT	67,518	,786	,377
Word Concretness in LT * Word Length in LT	10,954	,128	,722
Word Frequency in LT * Word Concretness in LT * Word Length in LT	76,188	,887	,348
Error	85,857		

Reading the table above we observe that only word frequency is a significant factor in test scores (both in SHT and LT tests). The other two properties do not appear to be significant in students' performance. In addition, we notice that word properties combined (as pairs and all three together) do not affect the scores. It becomes clear that, as far as our study in concerned, the role of word frequency, as measured by General Service List (GSL) (West, 1953) rather than actual (classroom) exposure, seems to be an important factor in learning new L2 vocabulary. The results presented here combined with the ANOVA results and Spearman Correlation (in sub-section 6.7.1) provide evidence for a positive effect of word frequency on test scores. Bearing in mind that unrelated vocabulary facilitates learning new L2 words (at beginners level), and since word frequency seems to assist L2 vocabulary learning, we could say that frequent words presented in an unrelated manner will promote ease of learning. This provides an answer to our **third research question** (see section 5.6). We may not be able to determine whether frequency is more important than unrelatedness (or the opposite) - because frequency is a word property and unrelatedness is a method of grouping for teaching presentation - however, they seem to be a powerful combination in teaching

new L2 vocabulary. Word length and word concreteness/abstractness, by contrast, seem to have no significant effect on test scores. Connecting this with the ANOVA results of sub-sections 6.7.2 and 6.7.3 we could say that only the unrelated presentation of L2 vocabulary facilitates learning in those cases. In conclusion, it is better to combine the unrelated presentation of L2 vocabulary with word frequency, while the combination with word length and word concreteness/abstractness makes no difference.

6.8 Qualitative data presentation in relation to quantitative analysis.

In this section, we complement our quantitative analysis with findings from the qualitative part of our research[52]. We will present qualitative data obtained from students' questionnaires and teachers' informal interviews. In the present study, we used qualitative data in order not only to contextualise the statistical results but also acknowledge the subjects' feelings and teachers' personal views and opinions.

6.8.1 Qualitative data and findings

Immediately following completion of the study, each student was asked to complete a post-test questionnaire in order to provide further information on students' views concerning the two alternative ways of presenting vocabulary (see Appendix 7). The subjects were invited to comment on the two different learning strategies used in the study. We used questionnaires to tap into the knowledge, opinions, ideas and experiences of our learners. Each student was asked which of the two vocabulary presentation strategies he/she found more helpful or more difficult and why. The students were also asked to comment on the use of word cards. It is essential to mention here that the qualitative findings below were only taken from the first sample of our study (intermediate students). Regarding adults students, we were not able to observe how vocabulary is introduced to the class because the permission for classroom observation was refused (due to the institution's regulations prohibiting the presence of any person other than the instructor). In addition, a post-test questionnaire was not administered to the adult participants due to time restrictions.

---

[52] According to Wallace (1998:38), there should not be any real opposition or contradiction between the two approaches. Quantitative data can throw light on qualitative insights and vice versa.

The most important findings from the analysis of the post-test questionnaires (see Appendix 7) are that most of the students (from both class A and class B) noticed the difference between the two ways of presenting new vocabulary by saying that in one way the words were organized having something in common (e.g. synonyms) and in the other the words had nothing in common. Interestingly enough, we notice the following paradox: while teenage students' scores in the tests were very low and the statistical analysis revealed that neither related nor unrelated vocabulary helped EFL intermediate students to achieve better scores, we observe that most of the students preferred related vocabulary to unrelated. Most school children (not adults) wrote that they found the related sets of words easier to learn and memorize (especially the topic-related words), because they were organized in a specific way having something in common. They also said that these words seemed easier to remember. This can be explained due to the fact that EFL coursebooks present vocabulary in groups such as the colours, foods and so on which students are expected to learn together (Waring, 1997). This idea is in accordance with the coursebook survey revealing that most of the coursebooks follow this pattern (see section 4.2.2). In this sense, the students are used to this kind of presentation showing their preference to semantically related vocabulary, which appears to contrast with the quantitative study.

Additional information acquired from the questionnaires (see Appendix 7) is that word cards helped students to learn new vocabulary items. They also stated that their favourite exercise was matching the English words with their Greek equivalents. This is to be expected, since writing new words and their translations in Greek was the commonest kind of note kept on new vocabulary, because it is seen as a simple bilingual rather than monolingual list (Scholfield and Gitsaki, 1996).

Besides the post-test questionnaire we had the opportunity for classroom observation. In order to monitor how vocabulary is introduced to the class we created a Vocabulary Observation List (see Appendix 6). We observed one class on 25/02/2004 (the teacher was an English native speaker) and one class on 27/02/2004 (the teacher was a non-native speaker). The purpose of classroom observations was to get an insight into vocabulary presentation in class, and also to make a record of teachers' views on vocabulary teaching. The findings, of course, are not fully representative as to how vocabulary is introduced in general because we observed only two classes. Bearing in

mind that teachers do not have much time to spend on vocabulary teaching exclusively during each lesson, we tried to collect as much information as possible. However, the findings give us some evidence as to how vocabulary is treated in a Greek EFL classroom.

The findings suggest that most of the time the meaning of new words is extracted from the passage or dialogue included in the unit. After the passage was read to the class, students were asked by the teacher if there were any unknown words. In that case, new vocabulary was mainly presented through context and explained through L2 paraphrase or synonym. Most of the vocabulary was dealt with explicitly (e.g. using L1 translation) but not in isolation (for example a new word was found in a grammar exercise). Vocabulary was generally introduced in topic-related sets, e.g. 'Great Inventions' as presented in the students' coursebook *Shine C* (Garton-Sprenger and Prowse, 2000:53). Also, in this coursebook there was no separate section of vocabulary because it was integrated along with the four skills (reading, writing, speaking and listening).

Concerning the procedures used to practise new vocabulary, the teachers used a variety of activities. For example, they asked students to practise vocabulary in writing (e.g. writing a paragraph using new vocabulary). The teacher also asked students questions in order for them to provide responses using new vocabulary (speaking). In addition, the students read the text in the unit which contained new vocabulary (reading).

Informal interviews were also conducted with the teachers after classroom observations and generated some very interesting findings in relation to vocabulary presentation. The teachers try to use English in the classroom as much as possible. They ask questions and give examples about new vocabulary in English. This means that they elicit the answer from the students by asking questions like, 'Do you know what *persuade* means?'. If the students do not know the word, the teacher tries to elicit the Greek equivalent using L2 synonym, antonym or paraphrase. If this is not successful, the teacher provides the Greek translation. This procedure is strongly supported by Scholfield and Gitsaki (1996). Teachers always try to highlight the new words on the board and they point out that they avoid explaining new vocabulary through the L2 paraphrase in order to gain time for grammar presentation. This makes the use of L1 translation very common and frequent. This finding is supported by Read (2004:146) stating that classroom

communication activities are centred almost entirely on the acquisition of grammar. Another interesting finding from teachers' interviews is that students do not spend much time studying vocabulary. They also claim that students spend most of their time in studying for Greek state school. This results in a lack of concentration and interest in L2 learning.

According to the teachers, presenting new vocabulary in topic-related sets helps students remember the words by making mental associations between these words. For example, if we say the word *beach* everyone will 'picture' it in their mind with the concepts of *sand*, *sun*, *scuba diving*, *sunbathing,* etc. They also point out that teaching topic-related vocabulary and synonyms works very well with advanced students. Another common characteristic of Greek EFL students is that while teachers ask questions in English, students tend to respond in Greek. It is also worth mentioning that students do not use dictionaries to look up new words, since they are provided with the Greek translation by the teachers.

6.9 Conclusion with summary of main findings

The present chapter dealt with the statistical analysis of the data obtained from the research described in Chapter Five. Statistical tests used for the analysis were reported and explained in detail. We found that:

- Adult beginners performed significantly better on the unrelated vocabulary test than on the related vocabulary test.
- Children (intermediate level) showed no significant difference in test scores between related and unrelated vocabulary.

This part extends the results of previous research which were only limited to beginners (Tinkham 1993, 1997, Waring 1997, Schneider, Healy and Bourne, 1998, Finkbeiner and Nicol, 2003). Regarding beginners, the result above is compatible with the results of previous research that semantically related vocabulary impedes learning while unrelated words seam to facilitate learning. It is crucial to mention that these results reinforce the positions stated by the researchers mentioned above, since they were extracted from a natural language in EFL classroom through teaching procedure.

We also noticed:

- the high scores of adult students in opposition to the low scores of children learners.

It seemed that adults were highly motivated and more conscientious learners for personal and professional reasons. They have also the advantage of being able to master certain aspects of a L2 even well into adulthood (Scovel, 1988). Children on the other hand, had less immediate and clear motivation for learning English vocabulary.

Another finding is that:

- adults performed better than children on synonyms.

A possible explanation for this is that L1 adult learners are able to extract semantic, syntactic and morphological information while becoming acquainted with the form of the word. This happens because there is a highly contextualised input. In addition:

- children tended to retain homonyms better than synonyms.

Phonology appears to play a more prominent organizing role in the L2 mental lexicon than semanticity. However, the number of words involved is too low to draw any definite conclusion from here. Another interesting finding is that:

- words to do with crime (topic 1) seemed to be better retained according to the test scores for both children and adults.

A possible explanation is that subjects are frequently exposed to crime-related vocabulary in movies with Greek subtitles.

Through a series of statistical tests we found that neither the length nor the concreteness/abstractness of a word seem to have any influence in test scores. Regarding word frequency, though, we found that:

- word frequency, when combined with unrelated presentation of new L2 vocabulary, makes a difference in students' performance.

The frequency of a word in language, however important for learning (as shown above), is also a usage factor dependent on the type of vocabulary presentation. According to Nation (2001), the positive role of word frequency in L2 vocabulary retention is to be expected. For this reason, Nation suggests that teachers and students should pay attention and spend considerable time on vocabulary.

The last part of the present chapter presents qualitative research based on interviews and questionnaires. Qualitative data and findings are displayed in connection to their role in overall discussion of the findings. This data revealed that:

- adults are highly motivated for personal and professional reasons
- children show lack of concentration and motivation because their time is occupied by studying for Greek state school
- vocabulary teaching is overshadowed by teaching of grammar

The next chapter (Chapter Seven) considers the main conclusions and provides an overall discussion on the different aspects that the present thesis has presented and developed. Further research questions will also be suggested.

CHAPTER SEVEN

Summary and Conclusions

7.1 Rationale and motivation for the present study

This study differs from similar ones in having been carried out in a natural setting. The use of a natural L2 combined with the teaching procedure in a real classroom environment makes this research generate results that might apply to natural L2 learners. On the contrary, previous research (Tinkham, 1993, 1997, Waring 1997 – see Chapter Four) was tightly controlled to benefit the researcher, not the learner, as Waring (1997:271) points out.

Waring (1997) refers to the nomothetic research tradition within his and Tinkham's (1993, 1997) work. The intention of quantitative or nomothetic research is to look for a single reality or truth. The qualitative or hermeneutic research tradition seeks to investigate learning in the classroom situation in naturalistic, interpretive or qualitative terms allowing for multiple realities (Ochsner 1979). Meara (1996:38-39) exemplifies the nomothetic view by saying that we need a "challenging combination of real-world constraints and rich theory", that is, we need a balance between the dominant nomothetic tradition and the hermeneutic tradition. According to Waring (1997), as the variables in his and Tinkham's studies were tightly controlled in those experiments, it renders them somewhat ungeneralizable for our classrooms, in other words, the results these experiments generated, might not fully apply to the natural environment of a classroom because they lack external validity. As Waring points out (1997:272), we need not only acceptable guidelines and agreed-on standards of measurement within the nomothetic tradition, but we also need guidelines for qualitative studies looking at the same aspects of language from a wider interpretive view within the hermeneutic tradition.

The present study tries to duplicate the real world application of the results found in the previous studies. The experimental design and the variables were not tightly controlled to benefit the researcher but the learner. It is clear that the present study points to a different approach concerning applications in a real L2 classroom. Though it is not

tightly controlled it promotes the use of natural L2 in natural settings. In this sense, we have a clearer understanding of the role of vocabulary presentation no matter what the drawbacks. The basis of this study was to examine whether the results of previous studies concerning the benefits of unrelated vocabulary presentation can be applied to L2 classrooms. We decided to follow an 'action research' model with elements of post-hoc research (statistical interpretation of test scores) inspired by replications of previous experiments trying to use real-world elements and circumstances (see Chapter Five). In order to achieve balance, this research tries to convey L2 vocabulary acquisition as part of a natural classroom environment by categorizing the presentation of vocabulary and examining word properties (see Chapter Five and Chapter Six).

7.2 Main findings

The results of the present study (see Chapter Six) support the view that presenting L2 beginner students with their new vocabulary grouped together in sets of syntactically and semantically similar new words may impede rather than facilitate the learning of the words. We have found that while adult beginners achieved significantly higher scores on the unrelated vocabulary test compared to their performance on the related vocabulary test, children (intermediate level) showed no significant difference in test scores. Regarding beginners, the result is compatible with the results of previous research, showing that semantically related vocabulary impedes learning while unrelated words are proven to facilitate learning (Tinkham 1993, 1997, Waring 1997, Schneider, Healy and Bourne, 1998, Finkbeiner and Nicol, 2003). Furthermore, the present results reinforce the positions stated by the researchers mentioned above because they were obtained by using a natural language in an EFL classroom through teaching procedure.

The results of this study indicate that adult beginners had better retention of new L2 words when those were presented in an unrelated fashion. The use of semantically related vocabulary is good for building networks in the mind (resembling the nature of the mental lexicon – see Chapter One), but not at initial stage. The children's results, however, being the intermediate group (in our study), and having a more established knowledge of the L2, do not demonstrate a positive effect of related vocabulary presentation. The effects found in this study may therefore be restricted to beginning

learners, rather than intermediate ones, as the beginning learner has to set up semantic and vocabulary knowledge networks in the L2 into which the words must be placed.

In addition to the above, we also noticed the high scores of adult students in relation to the low scores of children. Adult beginners are in some ways the easiest people to teach (Harmer, 1991:8), having a high degree of extrinsic motivation and succeeding very quickly. It seemed that adults were highly motivated and more conscientious learners for personal and professional reasons. Children, on the other hand, had no immediate and clear motivation because of immaturity and lack of interest.

Another important result was that word frequency seems to make a difference in students' test performance. According to Nation (2001), the positive role of word frequency in L2 vocabulary retention is to be expected. It seems that Corpora or Word Lists may prove to be useful after all. The effect of word frequency comes as an additional factor of L2 vocabulary acquisition. The frequency of a word in a language is also a usage factor dependent on the type of vocabulary presentation. One of the results suggested that word frequency when combined with unrelated vocabulary presentation may assist students in learning new L2 words.

From our results it appears that phonology may play a more prominent organizing role in the L2 mental lexicon (see Chapter One) than semanticity, since children in both short and long-term tests performed better in homonyms than synonyms. Also, adult learners seem to perform better in synonyms than children. This may happen because they are able to extract semantic information while becoming acquainted with the form of the word, since there is a highly contextualised input in their L1. It also seems that the exposure to crime-related vocabulary in movies with Greek subtitles may have affected test scores as we have seen that topic 1 (crime) seems to facilitate children's and adults' scores. In addition, qualitative findings present adults to be highly motivated. On the contrary, children do not have any immediate interest and motivation. Moreover, it was noticed that vocabulary teaching is overshadowed by teaching of grammar.

7.3 Limitations of the study

Finally, at this point it is important to mention limitations and possible contaminating factors in the study. It is limited in itself because there are practical constraints. Due to obstacles in obtaining official permission the size of the sample was small. It was difficult to acquire access from other schools to observe more classes and have a bigger sample. This has to do with the reluctance of public and private educational institutions to grant permission for research. The limited number of subjects leads to low generalizability.

There was also a limited amount of time because examination schedules reduced the time available for instruction. In addition, the adult educational programme lasted only for ten weeks. Only 60 words were taught and tested in each group (60 related and 60 unrelated). However, even though at first glance the number of words used in the study seems to be small, it appears adequate given that ten words per 1-hour session are regarded as a reasonable quantity to teach (Schmitt, 2000:144). Also, lack of interest, possible boredom and fatigue could be mentioned as potential obstacles to teaching a larger word sample.

Another limitation is the use of natural language with real EFL subjects. This creates the difficulty of isolating one variable from others due to lack of control. A language would need to be found where the learning for each word would be similar in terms of words with different stress patterns, the number of syllables (and so on). This would increase internal validity. Internal validity seeks to demonstrate that the findings (must) accurately describe the phenomena being researched (Cohen, Manion and Morrison (2000:107); in other words, it is concerned with the question of whether the study really did illuminate the effect of one variable upon another. It is difficult when teaching real, as opposed to artificial words, to isolate the variables of relatedness and unrelatedness from other variables that might affect retention of word meaning (e.g. frequency or length). The use of artificial language is seen as a positive feature because it increases the internal validity of the study by decreasing the possibility that particular artificial words might somehow be more or less learnable than others (Tinkham, 1997); in other words, artificial language is strictly controlled. However, this means that one cannot generalize the findings to natural languages without qualification being made (Waring,

1997:268). The drawbacks of using artificial language, however, include reducing the external validity and motivation of learners. External validity refers to the degree to which the results can be generalized to wider populations. The importance of the experimental results in terms of application to natural language learning is diminished through the use of artificial language. The question to be asked is whether artificial language creates conditions comparable with the 'real-world' situations or whether it encourages artificial responses in line with the artificial setting. As Waring (1997:269) points out, an attempt should be made to offset the artificiality of the words used in such experiments.

Another potential problem that needs to be addressed is whether the results of this study are representative. The issue here is that only a minority of research findings in the social sciences aspire to the 'law-like' status of many of those in the natural sciences such as physics (Wallace, 1998:44). It might be possible to prove statistically that a certain group-work task has had specified positive results. These cannot ensure, however, that it will have the same results in another country with different cultural norms; or even in the same country, in another context with different subjects. In this sense, most findings in an area of social sciences like education are indicative rather than truly conclusive.

We are conscious of the limitations. We cannot isolate all the factors that are combined and operate in a real EFL classroom. We realize that this study can only be suggestive and not decisive because there are confounding factors. However, this should not prevent the researcher from examining and investigating the effects in a naturalistic setting of presenting words in related and unrelated sets. These limitations should serve as clear indications that the study does not conclusively demonstrate the effects of semantic clustering but, rather, calls into question certain currents beliefs and instructional practices. An unavoidable interaction between properties of words in themselves and the way they are grouped for presentation creates possibility of confounding. All findings have to be interpreted cautiously. Nevertheless, this study is a worthwhile attempt to investigate natural settings using statistical methods in order to reduce lack of experimental control. Real world application is needed in order to benefit areas of linguistic interest (e.g. vocabulary learning and teaching). Since there is comparatively little research to report on methods of presenting and practising

vocabulary in the classroom (see Read 2004:153 for a short account of relevant studies) this study is considered to be a useful step forward.

7.4 Possible pedagogical implications

This research complements previous studies and suggests some re-evaluation of current pedagogical practice. The results of this study may come as a surprise to many current writers of ESL text who rely heavily upon the employment of semantic clusters in their presentation of new vocabulary. Consequently, L2 curriculum writers and programme planners who currently present students with clusters of semantically and syntactically similar new words may need to reconsider such practice. They may want to explore the possibility that they might ease the burden of L2 vocabulary learning by incorporating a rather semantically unrelated form of presentation (at least at beginner level).

As mentioned earlier, an intermediate (or more advanced) learner would probably already know many words from the semantic groups and when presented with new words may only need to add new words to an existing store, rather than create a new one from scratch. It may therefore be that activities grouping words with related meaning are best used at a secondary stage when the words can be recognized, some meanings have been acquired, and learners have reached a point where they will benefit from further opportunity to make connections and distinctions (Hedge, 2000:122-123). For initial presentation, we can present unrelated vocabulary and later at a more advanced level present semantically related vocabulary. We should present related vocabulary in a way that does not create an environment for interference effects.

But what can be done to minimize interference? Nation (2000) examines the question of how to minimize interference from three view-points: those of course-designers, teachers and learners.

1) West Frequency Counts of various lexical sets (e.g. *white*: 334, *blue*: 126, *pink*: 47) present a divergence in frequency of members of the same lexical set. This is often taken as evidence for the difficulty in using frequency as a vocabulary selection and frequency criterion. For this reason, it is difficult for course-designers, as well as teachers and learners, to appreciate that items in sets such as months and numbers are best learned, initially, when not learned together.

The criteria of usefulness (frequency or need) and avoidance of interference (ease of learning) are more important than aiming for early completeness of lexical sets (Nation, 2000:8).

2) Teachers should present the items at different times. That is, present the most useful of the items (according to frequency or need) first; then, after that item has been reasonably well established, introduce the next item. There is no research to tell us how well-established an item needs to be before it can be safely contrasted with its opposite, near-synonym or other members of its lexical set (Nation, 2000). These two items should be introduced at least several days apart. Teachers should also use widely-differing contexts. If *hot* and *cold* occur together in a course and it is difficult to present them at different times, then they should be presented in quite different contexts (i.e. *hot* can be used with collocates, such as *weather* and *water*, whereas *cold* can be used with collocates, such as *morning* and *meal* (Nation, 2000). Increasing the differences between the items will decrease the strength of the association between them, thus reducing the chances of interference. Waring (1997) also mentions that it might be advisable to mix these words into thematic rather than semantic arrangement instead. For example, *sweater*, *try on*, *cash register*, *striped,* etc.
3) Learners need to know about interference, how to avoid it, and what to do when it occurs. When using, for example, word cards to learn vocabulary, learners should keep similar items separated. When interference does occur (e.g. when a learner confuses *north* and *south*), the most effective way to deal with it is to find some mnemonic trick (i.e. the Keyword technique) to distinguish the items (Nation, 2000:9).

7.5 Future research

As we have mentioned in Chapter Four, even though theory states that related presentation of L2 vocabulary would have a positive affect on students with a more established knowledge of L2, the results of intermediate subjects are contradictory. Further research with intermediate and more advanced students seems to be necessary in order to clarify whether related vocabulary plays a prominent role in L2 learning at this level. It probably made no difference with our intermediate subjects because they were not motivated enough. It would be negligent to say that the results from intermediate

students provide a strong contradictory indication to the theory mentioned earlier- further research on this matter comes as a necessity to examine if theory stands correct.

Another aspect that might require further research is teaching and testing procedure. We need to consider whether we obtain the same results if both teaching and testing methods are different. For example, another study could teach and test items not in isolation but in context (exposing learners to vocabulary through reading). Waring points out that researchers have found that the occurrence of interference depends on the type of stimulus material. When meaningful passages are used rather than lists of words or nonsense syllables, no interference effects are found (Haberlandt, 1994:211). Keeping in mind that the present study used only nouns to examine the influence of vocabulary presentation, it would be interesting to observe the results if we taught and tested verbs, adjectives and other different parts of speech. Further research could provide the chance for more repetition (with word cards), to spend more time on vocabulary, and to promote the use of dictionary and autonomous learning (i.e. encourage independent learning strategies, as Schmitt (1997:255) recommends). It would be interesting to improvise a more natural teaching procedure to perform such experiments in natural settings. Chapter Two provides a variety of learning strategies and teaching methods that could be used in order to promote and improve the results of this study.

In addition, the productive dimension of L2 vocabulary learning in relation to the manner of vocabulary presentation could also be examined. The present testing was on the receptive use (seeing the L2 word and having to provide the L1 translation) of the words. Another study could test if the effect also occurred productively (subjects to be given the L1 word and to produce the L2 word). We used a definition recall test because we were interested in the form-meaning connection and we tested the receptive vocabulary knowledge by requiring L1 translations. The productive aspect could be checked by using writing and speaking tests or even by requiring L2 translations to see if the results contradict the Revised Hierchical Model (see 4.3.6).

It would also be interesting to see if the use of thematically related vocabulary in natural settings would provide us with similar results. According to Waring (1997), it might be advisable to mix these words into thematic rather than semantic arrangement in order to

minimize interference effect. It would be interesting to see if thematically related vocabulary presentation could be more effective than semantic or even unrelated vocabulary. Especially at beginner level, the present research along with previous studies found that a possible reason for semantically related words hindering learning of new L2 vocabulary is that they can cause interference effect. Thus, it is worth examining the actual impact of related vocabulary presentation by using thematically related lexis in order to exclude interference effect.

Regarding word frequency, the current study examined its influence on test scores as a side effect of the main research question of relatedness vs unrelatedness. Additional experiments, more oriented in word frequency as used in corpora and word lists, should be performed in order to examine if sources of frequency could become useful tools in L2 vocabulary acquisition. However, it is "the quality and frequency of the information processing activities (i.e. elaboration on aspects of a word's form and meaning, plus rehearsal) that determine retention of new information" (Hulstijn, 2001:275). According to Laufer and Hulstijn (2001), learners who wrote compositions using a set of target words remembered them better than those who encountered the words in a reading comprehension task, and learners who wrote the missing words in gaps in the reading text retained more of the words than those who just read marginal glosses (Laufer and Hulstijn, 2001). So, frequency in class as a measure of exposure to vocabulary could be more important than frequency in natural language. That is why Laufer (1997:141) mentions that "the frequency of a word's occurrence may be much different in a naturalistic, all-purpose language course as compared to a course in language for specific purposes". In conclusion, besides frequency in language, frequency in the classroom deserves attention, as it could prove to be a beneficial aspect of L2 vocabulary learning in a real classroom environment.

It is hoped that the present thesis has offered some evidence as to how the manner of organizing words for presentation may be important for learning new L2 words. It can also be considered as a useful ground for similar research as it presents certain standards of classification regarding important aspects of vocabulary learning and teaching (related, unrelated, frequent, infrequent, etc.). This though was only attempted in order to produce results in a natural L2 classroom by examining which factors seem to be helpful for the learner, not the researcher.