Durham E-Theses

Lexical input and categorization:: a study of vocabulary depth in second language learning

Liao, Fei-Hsuan

How to cite:

Liao, Fei-Hsuan (2003) *Lexical input and categorization:: a study of vocabulary depth in second language learning*, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/3741/

Use policy

Lexical Input and Categorization:

a study of vocabulary depth in Second Language Learning

by

Fei-Hsuan Liao

a thesis submitted for the degree of Doctor of Philosophy

School of Linguistics and Language

University of Durham

2003

Abstract

Lexical Input and Categorization:
a study of vocabulary depth in Second Language Learning

Fei-Hsuan Liao

This study investigates the association of vocabulary development and reading comprehension, focusing principally on the contribution of context to vocabulary depth learning.   Research into reading for vocabulary learning has typically emphasized the learning of discrete items as though they were unrelated.   Adopting the viewpoint of cognitive linguistics, this study assumes a cognitively motivated relationship inherent in these to-be-learned items, and thus defines vocabulary depth in terms of the ability to recognize the inclusion and membership properties of semantic categories occurring in a text.   In other words, what distinguishes this study from others is its emphasis on extent to which learners are able to recognize the hierarchical relations (inclusion) and internal category structure (membership) of the lexical items found in a text.

The research was conducted in a real classroom setting using procedures that are pedagogically valid.   This is believed to be the optimal context for revealing the genuine nature of second language learning.   Specifically, a Freshman English course incorporating extensive reading as one of its requirements was selected for this investigation since successful reading at advanced level depends upon recovering the instantial categories (i.e. those that exist in some particular instance) established in the text. To compare the effects of the original and alternate encoding options, a selected text was modified structurally and paratactically so as to enable a comparison of the extent to which text structure is associated with comprehension, when comprehension is defined in terms of the ability to recognize inclusion and membership.   In addition, the rhetorical properties present in the original text were also taken into account in order to investigate whether common rhetorical properties could result in more successful recognition of category membership and whether distinct rhetorical properties could lead to more successful differentiation of degrees of category membership.

The results revealed surprising findings. First, it was found that a more paratactic encoding could lead to more successful recognition of the inclusion properties of semantic categories. This finding provides significant information for materials development and especially for the type of reading materials used in this research, i.e. graded readers. Second, the presence of category names did not enhance the ability to recognize category members present in context, a result which contradicts the results found in experimental psychology. In addition, the effects of rhetorical properties on recognition of category membership *per se* and on differentiation of degrees of category membership were not parallel: while the ability to recognize category membership could possibly be enhanced where category members shared common rhetorical properties, the ability to differentiate degree of category membership was not found to be associated with distinct rhetorical properties among category members.

In an appendix to the thesis, the researcher also speculates on the possible effect of their native language conventional classifier system on the performance of the informants in recognizing instantial category membership.

Declaration

I confirm that this thesis is my own work and has not been previously submitted for any other academic award. Credit has been given to sources of reference that I have used.

Statement of Copyright

Acknowledgements

Table of Contents

List of Tables

Chapter 1   Introduction

1.1 General Background of Second Language Learning: the Input Issue

The discussion of input that follows will first consider a model of Second Language Acquisition (SLA) taken from the generative tradition. Subsequently, it will be shown that a cognitive linguistics approach provides a more appropriate basis for the current study, especially as it relates to one of the fundamental cognitive functions, categorization. In particular, we will show that a SLA model that originated in the generative tradition is especially convincing when operationalized within a cognitive linguistic framework.

Amongst the proposals which attempt to explain the processes by means of which language data received as input provides a basis for output, Gass' five-stage model (1988) is probably most frequently cited in the generative literature. According to Gass, five levels involved are 1) apperceived input, 2) comprehended input, 3) intake, 4) integration, and 5) output. In the first stage, while linguistic information reaches language learners, not all this information is recognized. Only those features noticed by language learners, i.e. the apperceived input, will go beyond their sensory perception. Several factors, including frequency of the occurring data, the effects of affective factors, prior knowledge and past experience, and degree of attention are likely to determine which aspects of language are noticed. These factors are liable to vary from learner to learner. Although *frequency* refers most obviously to the number of occasions on which features of input provided by the outside world occur and describes the typical characteristics that make the language data more available and thus enhance language acquisition, the quality of the input as well as the quantity also needs to be considered.

One paper reviews the theories concerning input quality by Saleemi, who presents three different approaches to the study of input, which consider input as triggering data, as comprehensible data, and as negotiable data (Saleemi, 1989: 178). The approach of input as triggering data mainly discusses the theory of Universal Grammar (UG) which is regarded as the initial stage of language acquisition, both for L1 and L2, whilst the approach taking input as comprehensible data is famously proposed by Krashen and his followers that the input available to the learner should be "comprehensible" in terms of the optimal linguistic level desired at $i + 1$ level, which is slightly beyond the learner's current knowledge. Input classified as negotiable data additionally emphasizes the interactional nature of the speech which can be regarded as a type of input. As a matter of fact, the last approach is intended to focus on the active participation of the learner. The interactive speech regarded as input is also mentioned in Pica's discussion (1991: 189). In addition, Alcon analyses the factors that can affect the input in becoming intake and makes a distinction between the input factors and the learner factors (1998: 347-348). The input factors are discussed in terms of the linguistic properties of the input, whether they are intrinsic, or are simplified, enhanced (input enhancement) or modified, and thus more available to the learner.

In addition, learner internal factors considered from a psychological perspective have the potential to influence language learning and act as facilitating or filtering force. Studies of affective factors referring to learners' attitudes, motivation, and language anxiety, for example, have obtained fruitful results in the last two decades. For instance, Gardner, 1980, 1988 and MacIntyre & Gardner, 1991, 1994, have shown that affective factors may have a positive or a negative effect and can facilitate or hinder learning. At the same time, cognitive psychology has revealed a

close relationship between attention and outcomes of subsequent learning. In the field of second language learning, there is strong evidence to support the claim that the power of attention can constitute an important threshold for input to be further analysed (Gass & Varonis, 1994; Wesche & Paribakht, 1999; William, 1999). Especially, the active participation of learners in the learning process is able to determine the outcome of learning (Swain, 1985). The effects of these internal factors were considered to be necessary for the input to be able to be processed by the learner (Ying, 1995: 185).

The second level, i.e. comprehended input, also accentuates the learner's role. Comprehended input first has to be distinguished from comprehensible input, as proposed by Krashen (1981, 1982, 1985) and widely accepted as prerequisite for language acquisition. In contrast to comprehensible input, which is controlled by input-givers, comprehended input is controlled by language learners and refers to the connection made between meaning and form (Gass, 1988). Thus emphasis is placed on the understanding process occurring in the learner's mental system, with precise levels which have not been fully explored and necessarily extend from the semantic to the syntactic. For Gass, however, though comprehended input is considered essential for input to become intake, it does not constitute intake.

Corresponding to this notion is a new concept of input enhancement, a revised version of consciousness-raising proposed by Sharwood Smith (1991, 1993). If the input accords with the learner's developmental stages, acquisition is mostly likely to happen. If it does not, it will have little or no effect. In other words, externally created input may not have the intended salience for the learner. That is to say, input has to be "enhanced" so as to ensure that its salience is increased relative to the developmental stage of the language learner, though external in nature. The

3

learner's readiness to perceive the properties of enhanced input is supposed to occupy a central place in language learning.

Similarly, accessible input is also defined in terms of the learner's linguistic development and focuses on the learner's readiness to acquire certain structures (Ying, 1995). Input can only be achieved through the learner's existing language knowledge, while comprehensible input, though essential for language acquisition, only focuses on the external manipulation, mainly by teachers or native speakers. The concept of accessible input is intended to resolve a situation in which the structures of the input are constructed far beyond the learner's current level of L2 knowledge and the learner comprehends the data via other cues, such as prior knowledge. The input data is said to be comprehensible but not accessible (Ying, 1995).

At the third level, the notion of intake has been conceived from either a process or a product perspective. A product perspective considers intake as part of input, while a process perspective regards intake as a stage in the process of integration into the learner's mental systems. Following Chaudron, (1985), intake is seen as a process rather than a product and is thus termed "selective processing" (Gass, 1997: 5), although Chaudron incorporates the several stages involved in to one single process (Chaudron, 1985: 2). More precisely, it is a process mediating between input and the ultimate L2 knowledge achieved (Gass, 1988: 206) and is closely related to the current linguistic level of the learner.

As a result, integration is the outcome of intake. Several cycles of analysis and reanalysis may be involved in the process of integration which brings about a development in L2 knowledge or storage of such knowledge. The impetus for developmental change is the mismatch discerned by the learner between their

existing knowledge and the incoming input (White, 1987; Sharwood Smith, 1991).

The learner's awareness of such a mismatch will arise either from the data,

especially when manipulated so as to make to-be-learned features salient, or from the

learner's current knowledge.

Finally, output provides the learner with opportunity to test their hypotheses

with regard to the target language (Swain, 1985), and has the crucial function of

pushing learners processing from the semantic level to the syntactic level in order to

provide comprehensible output.   This step enables the learner to establish

grammatical competence either by correcting inappropriate or mistaken hypotheses,

or by consolidating the existing system (Swain & Lapkin, 1994; Swain, 1995).

Nevertheless, as Gass argues, output is not equal to the grammatical system achieved

by language learners. Many factors can affect language performance, as noted by

Chomsky in his distinction between competence and performance (Chomsky, 1965).

In sum, the framework proposed by Gass covers important issues in language

learning in terms of five levels of operation. Although many of them remain

controversial or not fully explored, the framework provides not only a more

complete view of second language learning for applied linguists and researchers, but

is also an important guideline for language teachers and enables them to re-

conceptualise their classroom practice.

1.2 Motivation for the Present Study

While apperceived input, comprehended input, intake, integration, and output

are generally understood as important features in the second language learning, the

cognitive linguistics approach adopted in this study and specifically illustrated in

categorization theory provides a more tangible notion of apperceived input and

comprehended input. The examples provided in categorization theory are considered to be more concrete than the structures given in generative tradition. In addition, the cognitive approach makes a further attempt to investigate the more indeterminate process of intake and integration by drawing on our experience.

Unlike generative linguistics, which regards mental functioning as modular and thus autonomous, we agree more with the viewpoint of the Vygotskian approach that human mental functioning is always affected by the outer world and can never be autonomous. The influence of the social or sociocultural aspect is considered to occur first at the "intermental" or "interpsychological" level (i.e. between people) and then at "intramental" or "intrapsychological" (i.e. inside individuals) (Wertsch & Tulviste, 1996: 60-63). To understand the content or the nature of human mental functioning, a consideration and examination of social or sociocultural processes is necessary. The "social origins" of our mental functioning suggest that we need to look into the surroundings of individuals, including the society, the people and even the history of human beings, so as to seek for the shaping force of our cognitive structure. This draws our attention to the interaction with our environment and the people around us, and hence to our experience derived from these interactions. Therefore, it will be argued that a cognitive approach is more able to demonstrate that our conceptual system, including our linguistic system, reflects our life experience.

Following Vespoor and Winitz (1997), this research investigates the effect of lexical input on language learning, and especially on vocabulary learning in a second language context. We agree with Vespoor and Winitz's view that there is a practical problem in Krashen's proposal for comprehensible input (Krashen, 1981). That is, it is unlikely, or at least very difficult, to design language teaching materials according

to levels of increasing grammatical complexity. Therefore, whilst provision of comprehensible input at the level of $i + 1$ may be considered optimal, it is in fact not workable in practical teaching (Vespoor & Winitz, 1997: 63). Given this position, comprehensible input is defined in relation to the L2 lexicon. In other words, input for second language learners is considered in terms of lexical items and their properties, rather than the structural properties of language.

However, the lexical items that are treated as input are not randomly selected, but structured on the basis of lexical-field theory and semantic activation theory. In Vespoor & Winitz's study, subjects received lexical-field books for self-instruction (1997). The lexical items contained in these books were thematically related in stories or dialogues, as well as organised under superordinate terms, such as "transportation", or are semantically related items, such as "sliding" and "tumbles" in the "walk" book (1997: 66). It was assumed that the relatedness of lexical items would activate a semantic network that made easier the understanding and learning of a new item. Greater language improvement occurred in the group receiving lexical-field input than in the control group, as measured by a language proficiency test, the Michigan Language Battery. In effect, this research demonstrated the possibility that an alternative definition of language input that refers to the lexicon is workable and presumably holds promise for vocabulary learning, although it should be noted that the researchers did not specifically measure the vocabulary knowledge obtained by the learner.

Almost two decades earlier, the importance of the L2 lexicon was acknowledged by prominent applied linguists in second language learning because improper understanding of lexical items could result in more detrimental consequences than the inability to understand grammatical constructions (see Gass,

1989). Moreover, a surprising phenomenon for Second Language Acquisition is the lack of lexical theories, which renders SLA theorising incomplete (Meara, 1984; Gass, 1989). The lexicon and lexical theory have received comparatively less attention than syntactic theories and so far there are no detailed descriptions of the lexicon and no comprehensive lexical theories. As stated by Meara, the misconception that the lexicon is an "inherently messy part" of language is firmly rooted (Meara, 1984: 230). This belief is responsible for the relative paucity of lexical research.

Following Vespoor & Winitz's demonstration that arranging lexical items according to lexical field theory for input purposes has been shown to improve language learning in general (1997), we consider the likelihood that structuring lexical items in a specific way will improve vocabulary learning in particular. Since unlike grammatical rules, vocabulary knowledge is commonly regarded as unsystematic (Meara, 1984), to investigate the lexicon as a structured unit becomes a primary concern to correct this misconception. We suppose that the lexicon is organized in some way, but on account of the lack of research in this area, its actual nature is still unknown and remains speculative

The major approach of this research adopts a view of lexical knowledge which owes much to cognitive linguistics, a young field arising out of dissatisfaction with the contention of generative linguistics that the language faculty is modular and autonomous. In cognitive linguistics, language, like other cognitive abilities, is seen as a consequence of interacting with our experience of the world, and thus the language faculty is regarded non-modular. Linguistic structure, like general conceptual organization, largely depends on and is shaped by our world experience. The investigation of a language system is therefore able to reveal the "motivation

behind every aspect of language" (Dirven, Niemeier & Putz. 2001: xv). These underlying principles are valuable for second/foreign language learners because a foreign language becomes more tangible and therefore more learnable. Providing insight into the structure of a second/foreign language system helps the learner to find a rationale more comprehensible in constructing his own theory in relation to the target language.

While emphasizing general cognitive abilities, cognitive linguists identify categorization as a prime example of the connection between our perceptual experience and cognitive reasoning, both having been wrongly thought of as separate mechanisms (Lakoff & Johnson, 1999). Categorization is essential to our survival, and actually to the survival of all living things. We categorize objects and entities in the outside world, and our perceptual experience derived from our interaction with the material world determines how we categorize them. We perceive their similarities and differences with our visual faculty or differentiate their utilities based on their functions. That is, these categories are "part of our experience" (Lakoff & Johnson, 1999: 19). According to this argument, the basis of and the manner by which we categorize the outside world is determined by our neural structure, which can affect our perceptual capacity and the action we take to deal with objects and entities in the real world.

The close relationship of cognitive categorization and our experience is further shown by the evidence of basic-level primacy (Lakoff & Johnson, 1999; Rosch *et al.*, 1976). The distinctiveness of basic level categories reflects the ease with which category members can be recognized. This perceptual superiority of basic level over superordinate and subordinate levels is also revealed in four other aspects (Lakoff & Johnson, 1999: 27). First, the basic level is the level at which we perceive a distinct

figure. We can easily think of a concrete figure of a "chair", but we are not able to associate a similar concrete image with the superordinate term "furniture". Second, basic level category members have a similar general appearance, and thus a dog of whatever breed can be recognized as a dog. Third, we conduct similar actions towards objects at the basic level. No matter what kind of chair it is, we all sit on it; and no matter what kind of table, we all use the surface. Fourth, most of our knowledge is about the basic level rather than about the superordinate or subordinate levels. As earlier research has proved, significantly more features can be listed for basic level members (Rosch *et al.*, 1976). As such, while we conceptualise categories, the primacy of the basic level is so rooted in our perceptual experience that the formation of our reasoning and conceptual organization cannot be separated from our biological functioning (Lakoff & Johnson, 1999).

From the viewpoint of cognitive linguistics, linguistic structure is one part of our conceptual organisation and thus we find that every subsystem of linguistics presents the mechanism of categorization (Taylor, 1995). A simple and obvious example is that syntactically we categorize word classes as nouns and verbs in light of their similarity or difference with respect to structural behaviour. Not only can words be categorized into different classes, syntactic constructions are grouped into distinct types with various functions. Especially in accordance with the predictions of cognitive linguistics, the structure of these categories reveals prototype effects, family resemblance and fuzzy boundaries, just as has been discovered with conceptual categorization (Taylor, 1995). Therefore, the linguistic system is considered as one part of and inseparable from our general cognitive abilities shown by our propensity to categorize.

As linguistic categorization and general cognitive abilities are so closely related, categorization in a second or foreign language is assumed to have a crucial association with language learning. As envisaged by Dirven, "the description of such language-specific categories on the background of and in contrast to those available in the learner's native language" can provide a more insightful starting point for the learner to construct his own hypotheses in relation to the structure of the target language (Dirven, 1989: 57). Therefore, the cognitive linguistics account of differences and similarities between native language and target language can inspire the learner by drawing on his/her own real-world experience, thus creating a more profound understanding of how the target language organizes itself and how its linguistic units function. This understanding is believed to facilitate second and foreign language learning. Based on these assumptions, this study sets out to investigate the relationship between a cognitive mechanism and the learning of vocabulary knowledge in second language learning.

1.3 The Outline of This Thesis

There are seven chapters in this thesis. Each of the following chapters is summarised below so as to reveal the main ideas explored in this study and the connections between the various chapters.

Chapter 2 discusses empirical research studies in second language vocabulary learning. In the last decade, vocabulary learning has attracted more attention than even before, and a considerable amount of research efforts has been devoted to trying to discover effective methods of enhancing vocabulary learning for second language learners. In particular, given that reading has been found to contribute a great deal to vocabulary learning in the L1, L2 reading is seen as having the potential

to increase vocabulary breadth and depth by providing multiple exposures for L2 learners. In this sense, incidental vocabulary learning via reading is the primary focus of this research. This chapter will therefore explore the problematic issues selected to vocabulary learning from context. Notably, the danger inherent in this approach will be discussed and compensatory approaches will also be illustrated. In addition, the strategies adopted by the learner to deal with unknown or unfamiliar words occurring in a text will be considered and evaluated with respect to their effectiveness for vocabulary learning. Specifically, the strategies of lexical guessing, dictionary use and ignoring will be discussed and evaluated.

Vocabulary knowledge is generally agreed to cover two important aspects, vocabulary breadth and depth and chapter 2 accordingly summaries these two approaches. Vocabulary learning therefore involves acquiring knowledge along these two dimensions. Vocabulary breadth refers to the number of words learnt whereas depth is concerned with the extent to which knowledge of words is acquired. The distinction between these two aspects is significant both because this is the major conceptual distinction underlying recent approaches to L2 vocabulary research and because it is the depth of vocabulary knowledge that the current study investigates. In addition, some applied linguists distinguish the degree of control over vocabulary knowledge in terms of receptive and productive ability. The former is concerned with recognition of words and the latter with more active control over word knowledge, such as is required in writing. Since this distinction is emerging as an important aspect of vocabulary knowledge, it deserves some discussion.

In chapter 3 we first introduce the theoretical and philosophical assumptions of cognitive linguistics by means of contrasting cognitive linguistics and generative linguistics. The literature reviewed in this chapter explores the relationship between

word and concept, category and concept, categorization theory in cognitive psychology, and empirical studies involving word context and category context. In particular, in order to illustrate the effect of context on the determination of lexical meanings and on the internal structure of categories, a range of psycholinguistic experiments will be reviewed and discussed. After establishing the contextual effects with empirical findings, we introduce the underlying assumptions of the current research and explain the assumed relationship between vocabulary learning and reading comprehension which will be investigated in the study. A definition of vocabulary depth will be provided which rather emphasizes the hierarchical relations between lexical items occurring in a text. According to categorization theory, the internal structure of semantic categories is realized in inclusion and membership properties. Therefore, successful learning of vocabulary knowledge is defined in terms of the ability to recognize these properties as they occur in an extended text.

We describe the methodology of this research in chapter 4. First of all, five hypotheses in relation to language learning and the internal structure of semantic categories are proposed. These hypotheses relate specifically to two properties of category structure, inclusion and membership as identified in the previous chapter. Following the description and explanation of these hypotheses, the design of the experiment will be presented, including its setting, the pedagogic materials used as language input, the informants participating in the experiment and their second language learning background, and the instruments used for measuring the learning outcomes.

In chapter 5, we discuss the data obtained with respect to inclusion, and specifically the ability of learners to recognize category inclusion when the lexical prompts occur in both author constructed contexts and contexts modified by the

researcher. The results are reported and implications are drawn with respect to the development of language teaching materials, especially the writing of extended texts such as graded readers. In addition, we argue that vocabulary depth as defined in this research can provide an index of readability for publishers and writers to supplement the breadth-oriented indices currently employed. A consideration of the role of the classifier system in the mother tongue is also presented, and the possibility of the effects of the mother-tongue classifier system on the informants' ability to recognize the inclusion properties of instantial categories is discussed.

Chapter 6 presents an analysis of data obtained in relation to the ability of the informants to recognize the membership properties of semantic categories present in a text. In particular, this chapter reports on the association of the presence of category names and the ability to recognize category membership, and the association of rhetorical properties with the ability, to recognize category membership *per se*, and to differentiate degrees of category membership.

Chapter 7 discusses the implications of this study on the basis of the experimental findings reported in chapter 5 and 6. First of all, the limitations of the research are acknowledged, and the issues relating to obtaining data under classroom conditions, rather than more formal experimental conditions, are discussed. The chapter then discusses the relationship between vocabulary learning and reading comprehension, since, without successful understanding of the whole text, it would be difficult for learners to recognize the internal structure of the semantic categories that the author constructs. This also implies that vocabulary depth so defined can be taken as a criterion for checking readability. The chapter also discusses the place of vocabulary depth knowledge in vocabulary learning and teaching. Furthermore, as the study shows that the categorization behaviours of informants in dealing with

instantial categorization mostly involving events and actions are largely unexpected and different from the findings of existing experimental psychology research, it is speculated that the nature of the tasks used in language classrooms may play a role in categorization behaviours.

Chapter 2 Vocabulary Learning

In this chapter, I will review the major studies and empirical research relating to second language vocabulary learning. Although the importance of vocabulary in second language learning is acknowledged by earlier applied linguists, the investigation of vocabulary learning has been pursued more vigorously only recently. Therefore, the discussion will be set out with a brief introduction to the historical background of second language vocabulary studies, which indicates a comparative ignorance in this sub-field of linguistics and thus motivates the considerable number of recent research studies, amongst which the current study may be included.

Following a brief introduction in the second section, the relationship between context, reading and vocabulary learning will be examined so as to reveal whether or not they are as closely related as is assumed in most studies of vocabulary learning. Problems resulting from this assumption will be highlighted in order to clarify misconceptions about the effects of context. Possible solutions are also provided to reconcile the conflict between the advantages and disadvantages of vocabulary learning from context. In section three, reading strategies for dealing with unknown words will also be discussed. Three strategies frequently suggested in the literature, i.e. lexical guessing, dictionary use and ignoring, will be discussed with respect to their benefits and drawbacks. A combination of these strategies will be recommended for the optimal development of vocabulary knowledge.

In the fourth section, the components of vocabulary knowledge will be discussed in light of current theory. Notably, vocabulary breadth and depth are widely accepted as major facets of vocabulary learning. Vocabulary knowledge is therefore distinguished as to the number of words learnt (breadth) and the extent to

which knowledge of the way they are used is acquired (depth). In addition, receptive and productive knowledge are also distinguished, and the ease with which vocabulary knowledge can be retrieved, a subject of on-going debate. The discussion concludes by drawing on available empirical findings which explore the association of vocabulary knowledge and reading comprehension.

Finally, it is worth mentioning how this literature review motivates the current research. Although it is widely agreed that vocabulary knowledge and reading comprehension are closely related, how reading helps develop vocabulary knowledge both in breadth and depth has not been demonstrated empirically. This provides an important starting point for the current study, which is intended to reveal that reading can provide a principled context for the learner to acquire vocabulary knowledge, and especially vocabulary depth knowledge.

2.1 A Brief Historical Background

It is generally agreed that the time has come to put more emphasis on teaching vocabulary. Indeed, as early as 1976, Richards pointed out the neglect of vocabulary studies and called for more attention to be paid to vocabulary teaching and learning. Language teachers frequently found that learners expressed the worry that they did not have sufficient vocabulary knowledge to pursue further study, or that a small vocabulary size caused enormous difficulties in understanding texts. However, unlike syntax, which is systematically rule-governed, vocabulary knowledge is typically seen as unsystematic, a feature especially noted by Meara (1984: 230). Saussurean "arbitrariness" is typically understood to refer to the apparently random association of a phonological matrix with semantic representation. This apparent lack of systematicity in the lexicon has been considered an obstacle to providing

practical information for language teaching. Whether the reasons for the long-term

neglect of vocabulary learning and teaching were attributed to trends in linguistic

theory (Richards, 1976: 77) or the unique nature of the lexical system (Meara, 1984),

the impact of vocabulary knowledge on successful language learning is no less than

that of syntactical knowledge, as Gass has explicitly stated (1989: 92).

Apart from the lack of practical information that might be useful to language

teachers, another unbalanced theoretical development with respect to language input

for L2 learners was that, while research focusing on syntactic or interactional

structures has revealed valuable findings, until relatively recently few research

findings have been available with respect to vocabulary learning and teaching.

Nevertheless, increasingly interest has been focused on this area and has resulted in a

considerable number of research studies of importance (Bensoussan & Laufer, 1984;

Crow & Quigley, 1985; Li, 1988; Davis, 1989; Mondria & Wit-de Boer, 1991;

Hulstijn, 1992; Huckin & Bloch, 1993; Hulstijn, 1993; Luppescu & Day, 1993;

Jacobs, Dufon & Hong, 1994; Knight, 1994; Ellis, 1995; Hulstijn, Hollander &

Greidanus, 1996; Lawson & Hogben, 1996; Prince, 1996; Grabe & Stoller, 1997;

Paribakht & Wesche, 1997; Watanabe, 1997; Tinkham, 1997; Zimmerman, 1997;

Grace, 1998; Fraser,. 1999; Nation & Wang, 1999; Paribakht & Wesche, 1999;

Nation, 2000; Schmitt & Cater, 2000; Wesche & Paribakht, 2000; Zahar, Cobb &

Spada, 2001). In their study of lexical input, Vespoor and Winitz's (1997) even

suggest that lexical input be taken as an alternative way to provide appropriate

language input for learners. In addition, in the last two decades, there has been more

and more interest in how context can help learners with vocabulary learning and,

drawing on work in L1 reading research, whether incidental learning can bring about

vocabulary learning. Given these developments, it is the purpose of this chapter to

review the following relevant issues: the relationship between context, reading and vocabulary learning, the reading strategies employed in reading and their relationship to vocabulary learning, and, finally, aspects of vocabulary knowledge that have been widely discussed.

2.2 Context, Reading, and Vocabulary Learning

2.2.1 Context and Vocabulary Learning

Intuitively, many language teachers consider context as crucial for vocabulary learning because of the multiple exposures to word meanings it provides for language learners, thus revealing the specific properties of lexical items and the fine distinctions between similar words. Basically these assumptions are derived from L1 reading research (Nagy, Herman, & Anderson, 1985; Nagy & Herman, 1987; Sternberg & Powell, 1983; Dubin & Olshtain, 1993). More notably, the research of Nagy et al. (1985) has demonstrated that word learning from context does occur, though to a limited degree, and that, in contrast to direct vocabulary instruction, for long-term retention vocabulary learning through context is more effective than direct vocabulary instruction.

How can context help vocabulary learning? In this thesis, we are mainly concerned with the written context[1]. A theory of learning word meanings from context in L1 was proposed by Sternberg and Powell (1983). They identified possible contextual clues that would help readers infer word meanings. For example, temporal clues, such as whether the action described by the unknown word was frequent or whether the time at which the action occurred assisted the reader in guessing its meaning; spatial clues, such as the position of an unknown word; and

---

[1] Note that in L1 research, the oral context is equally important for vocabulary learning.

functional clues which reflected the purpose of an unknown word (Sternberg & Powell, 1983: 882). These clues might not all exist together, but readers could use one or more to infer the meanings of unknown words. Accordingly, in L1 reading research, vocabulary learning and reading are seen as reciprocally beneficial (Sternberg, 1987: 90). Learning new words is enhanced by contextual reinforcement and reading will be improved if more words are learnt.

Moreover, what is more important is that the linguistic environment in which unknown lexical items are embedded will encourage language learners to be engaged in more mental work by attempting to guess the meanings of unknown words (Scholfield, 1997: 295). It has been famously claimed that reading is a "psycholinguistic guessing game" (Goodman, 1976: 260) in which readers are supposed to use, or to learn to use, appropriate clues to guess the message by virtue of an interaction between thought and language. In cognitive psychology, this kind of mental effort involves depth of processing that has been proved to be decisive for long-term retention because a more enduring memory trace is created (Craik & Lockhart, 1972: 675).

More reasons in favour of presenting words in context are spelt out by Schouten-van Parreren (1989), who drew attention to four drawbacks of learning words in isolation. First, a word presented in isolation provides no "cognitive hold", or mechanism able to sustain a trace in learners' memory, and thus is liable to be forgotten (1989: 76). Secondly, isolated words are very often organized via thematic or alphabetic features, which were likely to result in interference which may be detrimental for vocabulary development. Thirdly, a meaningful word is always defined by its context, and therefore words in isolation are relatively meaningless. Finally, since words in isolation do not constitute a meaningful context, they cannot

arouse any affective engagement from learners, which is considered to be influential for retention. In terms of memory trace, context provides more opportunities for a learner to form connections between a previous trace transformed from learner's experience and new words. The possibility of retention of new words thus becomes greater (Schouten-van Parreren, 1989).

Based on these assumptions, many applied linguists and language teachers suggest that reading is definitely beneficial for building L2 vocabulary, with the strongest position taken by Krashen (1989), who claims that extensive reading is necessary for vocabulary learning. The merit of extensive reading lies in the opportunity it provides for multiple exposures, which have been found to be insufficient in other L2 learning contexts (Nuttall, 1982; Hafiz & Tudor, 1989; Klapper 1993; Coady, 1997). Decontextualized vocabulary instruction, on the other hand, cannot provide enough exposures for vocabulary learning, or sufficient opportunities for learners to make sense of word meanings within context. In Nuttall's words, "an extensive reading programme is the single most effective way of improving both vocabulary and reading skills in general" (Nuttall, 1982: 65).

Indeed, empirical studies confirm the positive role of context and support the claim that context can enhance successful guessing and may be conducive to vocabulary learning (Li, 1988; Hulstijn, 1992; Huckin & Bloch, 1993; Grabe & Stoller, 1997). In their study of three individual L2 learners, Huckin and Bloch found that 23 out of 25 instances of successful guessing could be attributed to immediate context clues that occurred in the same sentence, especially collocating words. Furthermore, Li demonstrated that, for both reading and listening modes, the groups receiving cue-adequate information inferred word meanings with greater ease and gained higher scores for word retention than the groups receiving cue-inadequate

information. In Hulstijn's research, learners who engaged in more mental effort while inferring word meanings (as required to discriminate marginal glosses formatted as multiple-choice options) remembered word forms and meanings better (Hulstijn, 1992).

Nevertheless, vocabulary learning from context cannot be said to be absolutely beneficial. Some potential problems exist that merit a closer scrutiny. I will return to this topic in section 2.2.3.

2.2.2 Incidental Vocabulary Learning and Vocabulary Instruction

From the above discussion, linguistic context is known to be particularly helpful for vocabulary learning. One important assumption stems from the result of L1 vocabulary learning research that describes L1 vocabulary acquisition as "incremental" and "gradual" (Nagy et al., 1985: 236). This raises the notion of incidental learning. It is claimed that L2 learners, like L1 learners, will gradually acquire vocabulary during the process of reading without any explicit intention. "Incidental" refers to the manner of learning, i.e. without explicit attention paid to lexical input. Incidental vocabulary learning can also occur via other modes, such as oral input, television subtitles or other media. Incidental vocabulary learning by reading, however, has received a lot of attention in second language vocabulary learning owning to the influence of L1 research. Such a view of vocabulary learning has informed most of the recent published research studies (e.g., Hulstijn, 1992; Hulstijn et al., 1996; Paribakht & Wesche, 1997; Huckin & Coady; 1999), with a special edition of *Studies in Second Language Acquisition* (vol. 21, no. 2, 1999) dedicated to this issue. Nevertheless, a thorough discussion is necessary in order to eliminate misconceptions.

First of all, we have to define "incidental vocabulary learning". According to Huckin & Coady (1999) in their recent review of vocabulary learning, "incidental" means that no specific attention is paid to the vocabulary in question when learners are engaged in reading (1999: 184). In other words, the focus of attention is not on words, but on the overall meaning of a text (Gass, 1999: 320, Nation & Wang, 1999: 360). Furthermore, Paribakht & Wesche remark that vocabulary learning is incidental because it is seen as a "by-product" of reading comprehension (1999: 197). Even so, Gass argues that the definition of incidental learning is questionable since it is impossible to show that the learning is genuinely incidental and occurs without any attention from language learners (1999: 320), and Read comments that "incidental" should not be taken to mean "unconscious" learning (Read, 2000: 44). In principle, any kind of successful learning requires some degree of attention (Schmidt, 1990; Huckin & Coady, 1999). Gass further suggests that the term that characterizes the essence of L2 vocabulary learning should be "gradual" or "incremental" rather than "incidental" (Gass, 1999: 319). Using the word "incidental" actually neglects the role of learners, who must give some attention to individual vocabulary items in spite of the fact that the concentration is on the message or the overall meaning of a text.

An additional logical problem in taking results from L1 research and applying them to L2 learning is the possibility that, although incidental vocabulary learning underlies L1 vocabulary acquisition, this does not necessarily imply that L2 vocabulary development takes the same route or that incidental learning from context can account for L2 vocabulary development. In support of this view, Haynes points out that the fundamental difference between L1 and L2 vocabulary learning is the much larger quantity of vocabulary experienced and used by L1 learners,

compared to the limited repertoire of L2 learners (Haynes, 1993). Some scholars claim that a threshold vocabulary needs to be explicitly taught to L2 learners, and especially to beginners. Otherwise, reading becomes more difficult (Nation, 1990; Nation & Waring, 1997). Coady observes this problem and terms it "the beginner's paradox", a way of referring to the situation facing language learners with limited vocabulary knowledge who cannot therefore read well, and cannot benefit from the provision of context (Coady, 1997: 229). This situation is aggravated when the failure of L2 reading leads to frustration and distress, which in turn affects the inclination of L2 learners to read more, resulting in a "vicious circle" (Nuttall, 1982: 167). A similar problem can even occur with higher-level learners when they read authentic texts in which too many new words are embedded (Dubin, 1989; Coady, 1997). Therefore, vocabulary instruction is necessary, and is held to be particularly useful, and indeed indispensable in the initial stages of vocabulary learning (Nation, 1990; Oxford & Scarcella, 1994; Schmitt, 2000).

Typically, semantic field theory is used to structure lexical items for instruction. Simply stated, lexical items are arranged according to their semantic relationship rather than being randomly, or alphabetically structured (Channell, 1981). From this, we can see that vocabulary can be treated as systematic, rather than randomly structured. Crow and Quigley's study, for instance, demonstrated the superiority of the semantic field approach over the traditional approach, especially for long-term retention (1985). Subjects read passages and were supposed to learn unknown or unfamiliar words in the texts. The experimental group received a semantic field approach, which first identified words to be learned for each passage, selected a key word and four other related words for each target word and thus formed a keyword group. Exercises were given, emphasizing the semantic association within keyword

groups, for example, by replacing keywords with related words. The control group, on the other hand, received a traditional approach, consisting of word lists with definitions and exercises including word-level and sentence–level analysis, matching and substitution. Although the immediate post-instruction tests showed advantages with the traditional approach, a follow-up test found the semantic field approach to be superior (Crow & Quigley, 1985).

However, more recent research has also found that semantically categorized lexical items can cause interference (Tinkham, 1997). In this study, artificial lexical items paired with English words were arranged in clusters of semantic vs. unrelated items, and thematic vs. unassociated items, and were presented to learners. Items in the semantic cluster were of the same form-class and selected from the same superordinate concepts, while words in the unrelated clusters were not from the same superordinate concept. Thematic clusters contained words of different form-class but were thematically related, while unassociated clusters consisted of words of different form-class and did not share the same thematic concept. In contrast to previous findings, words in semantic clusters were found to cause interference, and new items were more likely to be retained in the case of unrelated word group exposure. However, words in thematic clusters were learned better than in unassociated clusters. Subsequently, Waring replicated this research with Japanese subjects, and paired Japanese L1 words with artificial words (Waring, 1997). The same result was found: semantically related words were more difficult to learn than unrelated words. Interviews were conducted with the subjects, which supported this result and revealed the difficulty of learning words with the same superordinates.

Since this way of organizing lexical items for the learner is prevalent in most coursebooks, Nation provided a practical suggestion to avoid the interference effect.

That is, by providing a normal context for these items to enable the learner to discriminate their usage (Nation, 2000). For example, different collocating words can be used to distinguish *hot* and *cold*, while the former can be presented with *weather*, *water* and *summer* and the latter with *morning*, *meal*, and *drink* (Nation, 2000: 9). In this case, context contributes to word meanings in a practical way.

Among the empirical studies on vocabulary instruction in second language or foreign language learning, one mnemonic technique, the keyword method, deserves mention. Two stages are involved in this method. The first stage is to find a keyword (a meaningful word), which is phonetically similar to part of or to a whole target word, and the second stage is to link the meaning of the keyword to the meaning of the target word with an imaginary picture. Atkinson called the first stage an "acoustic link" and the second stage an "imagery link" (Atkinson, 1975: 821). For example, to learn the Spanish word *carta*, which means *letter*, the keyword *cart* can be selected because of its similar sound. Then the image of a letter put inside a cart is formed to connect the meanings of the keyword and the target word (Pressley, Levin, Hall, Miller & Berry, 1980: 163). In second language or foreign language learning, the keyword is always a word from the L1 based on the assumption that the phonetic or orthographic familiarity of L1 words is able to enhance recall.

One of the earliest studies of the keyword method was conducted by Atkinson and Raugh (1975). Native speakers of English learning Russian were assigned to a control group (i.e. no particular strategies required) and a keyword group and were tested on three consecutive days immediately after exposure and also after a 30- and 60-day delay. All the results showed that the subjects in the keyword group performed significantly better than those in the control group. A recent study also showed that the subjects learning English vocabulary with Spanish keywords also

scored significantly higher than a control group, both on cued recall and sentence completion. The effect was also shown to hold after a one-week delay (Avila & Sadoski, 1996).

After empirical experiments found that groups exposed to the keyword method were superior to no-strategy groups or to groups using rote rehearsal strategies (Pressley *et al*, 1980; Pressley, Levin, Nakamura, Hope, Bispo & Toye, 1980), it seemed appropriate as a next step to investigate whether the keyword method was really superior to contextual learning. In an early experiment which attempted to compare the effectiveness of the keyword method and contextual learning, McDaniel and Pressley (1984) randomly assigned subjects to four conditions: no-strategy control, keyword, context, and combined keyword and context. The control group was provided with the definitions of target words, the keyword group with keywords and the definitions of the words, the context group with three sentences in which target words were embedded but without definitions, and the combined group with keywords and three sentences of this kind, i.e. also without definitions. A recall test was conducted, requiring the subjects to write a short definition of tested words. It was found that the keyword method was much more effective than context and combined conditions, though not more effective than the control condition.

Nevertheless, a different study presented an alternative finding (Moore & Surber, 1992). For subjects learning a foreign language (German), they found no significant difference between the scores for the keyword group and the context group, though both groups were superior to the control group. Although this result seemed to contradict the earlier study by McDaniel and Pressley (1984), one noteworthy aspect was that the previous study involved native English speakers studying rare English words while the latter study was concerned with foreign

language learning by subjects without a German background and thus might be more applicable to second language learning.

Using an experimental design similar to McDaniel & Pressley (1984), different findings were obtained with Spanish subjects learning English (Rodriguez & Sadoski, 2000). Four learning conditions were created, including rote rehearsal, keyword, context and context/keyword. Subjects in the condition of rote rehearsal were given English words and Spanish equivalents, those in the keyword group were given Spanish keywords and Spanish equivalents, while in the context group, subjects received three sentences containing the same target word whose meaning they had to infer and subsequently were given the correct meaning and tried to remember the meaning. As for the context/keyword group, subjects were also given three sentences containing a target word whose meaning they were required to infer but were provided a keyword to assist them in remember its meaning. Apart from the addition of the verification stage, a major difference from the previous research was that in this research subjects were either assigned to an immediate recall or a one-week delayed recall test. The superiority of the context/keyword group was confirmed, at a level of statistical significance. The value of the combined context/keyword method for vocabulary learning was therefore highly recommended (Rodriguez & Sadoski, 2000).

To sum up, the keyword method has been empirically proved to be effective in vocabulary learning, but has not surpassed the effectiveness of learning from context. Moreover, it has not been recommended as a replacement for other approaches, but only for fast retrieval as a memory task (Cohen, 1987: 59) or as a complementary method to contextual learning (Hulstijn, 1997a: 220). Even the strong advocates of the keyword method agree with this view of its value (Pressley, Levin & McDaniel,

1987: 123). Despite the good results obtained in keyword studies, learning from context should be considered the most natural and essential method for vocabulary learning.

From the research discussed up to this point, it is not possible to determine whether incidental vocabulary learning (learning from context) is superior to explicit vocabulary instruction, or *vice versa*. Therefore, as has been mentioned, a synthesized approach which takes explicit vocabulary instruction and incidental vocabulary learning as complementary seems better (Coady, 1997: 235; Schmitt, 2000: 120). Paribakht and Wesche attempted to empirically reconcile explicit vocabulary instruction and contextual effects by means of interventional instruction (1997). In their experiment, two groups were established, one designated as Reading Plus (RP) and the other as Reading Only (RO). The former received both reading texts and vocabulary exercises while the latter received reading and supplementary texts. In the case of the RP group, special attention was drawn to the linguistic characteristics of the target words in the vocabulary exercise. The results showed that the RP group achieved higher vocabulary gains than the RO group (Paribakht & Wesche, 1997). In addition, Zimmerman (1997) reported a study which combined the effects of reading and vocabulary instruction, with the experimental group receiving extra vocabulary instruction and as a result gaining higher mean scores on the post-test. Whilst emphasising the positive effects of reading, Zimmerman, therefore, encourages language teachers to spend some time on vocabulary instruction as well, rather than assuming that vocabulary learning will occur automatically during the reading process (Zimmerman, 1997). More specific and practical suggestions come from Nation (1990). Whichever approach is adopted, it should vary with different types of words and different proficiency levels of learners.

For high-frequency words, explicit vocabulary instruction may be necessary for low-level learners or for beginners, whilst for later stage learners a context-based approach may be more beneficial.

2.2.3 Context and Vocabulary Learning—problematic issues

First and foremost, reading for comprehension and reading for vocabulary learning are not the same thing and might even compete with each other in terms of cognitive processing. It has been shown that learners are almost always able to comprehend the provided message but are often unable to remember word meanings which have been successfully guessed or even verified by dictionary definition during reading (Li, 1988; Hulstijn, 1992; Grabe & Stoller, 1997). In their review of the relationship between contextual guessing and vocabulary learning, Nation and Coady raised the issue of the relationship between guessing and word retention (Nation & Coady, 1988: 101). Ellis' findings also revealed that no retention took place after learners had understood the message (Ellis, 1995: 425). The two subjects in Parry's study (1997) presented a contrast in their performance: one comprehended more but retained less; the other comprehended less but retained more. How then are we to account for these apparently paradoxical findings?

As long ago as 1972, it was suggested that the retention of vocabulary might involve some kind of deep processing that creates deeper traces in memory (Craik & Lockhart, 1972: 675). Based on self-report, Lawson and Hogben, for example, investigated various strategies that learners used to memorize target words (Lawson & Hogben, 1996). High correlations were found between elaborative strategies and retrieval of target words. Elaborative strategies included sentence translation, simple use of context (simple guessing from context clues), appearance similarity, sound

link, or more complex ones such as a complex use of context (referring to other words in the context), paraphrase and mnemonic. They all involved an attempt to generate the meaning of the target words by referring to the context or activating existing knowledge, and thus created deeper traces in memory. Therefore, acquisition of word meanings from context needed to be distinguished from comprehension of word meanings in context in that the former might involve more elaborative strategies as described above. Nevertheless, no noticeable result was found for association between the use of context and the retention of word meanings (Lawson & Hogben, 1996).

Secondly, context is not necessarily beneficial for vocabulary learning because context is not guaranteed to provide appropriate contextual clues for correct guessing (Dubin & Olshtain, 1993; Haynes, 1993; Laufer, 1997), and because the learner may just ignore these clues (Bensoussan & Laufer, 1984; Lawson and Hogben, 1996; Paribakht & Wesche, 1999). Haynes (1993) was especially doubtful about the claim that since L1 readers guess word meanings from context, L2 readers might also benefit from such an approach. In particular, he drew attention to the fundamental difference between L1 and L2 vocabulary learning and the large amount of lexical input encountered and used by L1 readers compared to that encountered by L2 learners (Haynes, 1993: 58). A major flaw of the contextual guessing hypothesis is the inability of context to provide enough clues for language learners, and therefore he suggested that other strategies needed to be developed to compensate for this limitation. (*ibid.* 55).

Nevertheless, it is found that the extent to which context is beneficial for vocabulary learning varies (Nation & Coady, 1988; Dubin & Olshtain, 1993; Haynes, 1993). Given the results obtained from an L1 cloze test, Dubin & Olshtain showed

that lexical items could be classified as providing high, mid, and low textual support in terms of the number of exact or close substitutes provided by informants as a function of the degree of sufficiency of contextual clues. They collected lexical items that were correctly answered or related to the target items. Words of high textual support were those that could be retrieved exactly or be supplied with equivalent substitutes, whilst words of low textual support were those that could not be either recovered or replaced by close equivalents, indicating that the text imposed insufficient constraints (Dubin & Olshtain, 1993: 191). This problem was also highlighted by Laufer, who regarded some contextual clues as "nonexistent", "unusable", "misleading" and "partial" (Laufer, 1997: 28-9). As indicated, these clues were not found as useful as they should have been. It was therefore suggested that guessing strategies can result in serious problems and that it is not sufficient solely to depend on the learner's guessing strategies to deal with unknown words.

A third problem is that encouraging contextual guessing will greatly impair vocabulary development if incorrect meanings are guessed and then retained. Modria & Wit-de Boer's research (1991) confirmed such a possibility. Their subjects retained a slightly higher percentage of incorrectly guessed words than of correctly guessed words. For this reason, Mondria and Wit-de Boer suggested a learning stage during which the learners are instructed to eliminate incorrect meanings (1991: 257). In other words, verification of inferred word meanings should be considered important in vocabulary development if contextual guessing is promoted in language learning.

A fourth issue is that context is not particularly helpful for learners of low proficiency who are not able to access sufficient syntactic knowledge to benefit from the contextual environment in which a lexical item occurs (Nation, 1990; Prince,

1996). Indeed, this is the main reason why both Laufer and Nation argue for the

establishment of a threshold vocabulary for beginners (Nation, 1990; Laufer, 1997).

Additional help might be needed to achieve reading comprehension, such as

providing an L1 translation for low-proficiency learners (Prince, 1996) and

encouraging dictionary use (Knight, 1994).   In Prince's study, subjects were

presented either with context in which words to be learned occurred in sentences, or

with translation pairing English words with French equivalents. It was found that L1

translation was particularly effective for low-level students because context required

syntactic decoding or analysis and resulted in great difficulties for low-proficiency

learners. Learners of low verbal ability were also found to benefit more from the

provision of dictionaries relative to those of higher verbal ability (Knight, 1994).

Their performance was better than that of the group working without the assistance

of dictionaries.

Fifthly, some research suggests that, apart from the issue of whether context

can help vocabulary learning, the provision of glosses should be also considered as a

means of providing word meanings which can be accessed by learners. A study by

Davis showed that glossing could facilitate the reading process though no

measurement of recall of vocabulary was conducted, with only the recall of reading

passages being measured (Davis, 1989).   Subsequent studies of the effects of

marginal glossing found positive results (Hulstijn, 1992; Hulstijn *et at.*, 1996;

Watanabe, 1997). Hulstijn, for example, compared vocabulary learning by a group

provided with multiple-choice glosses and a group with L1 translation equivalents

during the process of reading for comprehension. A higher retention rate was found

for the former group (Hulstijn, 1992). Watanabe modified context by providing

simple marginal glosses, multiple-choice marginal glosses, or appositives to target

items as experimental conditions (Watanabe, 1997). The results of vocabulary tests, both an immediate post-test and two delayed tests, showed that the groups provided with simple marginal glosses and multiple-choice marginal glosses scored significantly higher than the other groups. Furthermore, Hulstijn *et al.* (1996) showed the provision of marginal glosses to be superior to dictionary use in incidental vocabulary learning.

From these studies, a very important additional finding emerged, namely that the multiple-choice format of marginal glosses requiring learners to make a choice between possible lexical items on the basis of contextual information actually engaged them in more mental work and hence led to stronger retention of target items. Two aims can be achieved by this procedure: one is to avoid wild guessing and the other is to engage learners in mental work. Not only does the procedure encourage contextual guessing, it also illustrates how guided guessing can be introduced into pedagogic materials. However, Hulstijn warned that such procedures should not be introduced without the teacher's assistance to check that incorrect inferences are not occurring (Hulstijn, 1992: 123).

Jacobs, Dufon and Hong conducted a study on the effects of glosses on vocabulary learning involving issues such as word retention for both short-term and long-term periods, the kind of preferred language gloss (i.e. in the target language or the L1), whether or not this preference was related to vocabulary learning, and the favoured location for the glosses (Jacobs *et al.*, 1994). Since one merit of glosses is the provision of an opportunity to verify inferred meanings, which is considered essential for vocabulary development, it is important to know whether the learner reacts positively to them. Their findings revealed that overwhelmingly most subjects (99%) prefer glossing, that they preferred L2 glosses provided that these glosses

were understandable, and that marginal glosses were the type preferred by almost all subjects.

So far, the effects of context for reading comprehension and for vocabulary learning have been thoroughly discussed, covering benefits and possible dangers. For effective vocabulary learning and teaching, language teachers should be aware of these factors.

2.3 Reading Strategies And Vocabulary Learning

The learning strategies employed by second language learners have been an important research field ever since we acknowledged learners as autonomous individuals in language learning. To overcome unknown vocabulary presented in a text, L2 learners are known to use various reading strategies to achieve comprehension. The literature identifies three major strategies which are widely used during the reading process: lexical guessing, dictionary use and ignoring.

2.3.1 Lexical Guessing

Guessing has been regarded as an important skill in language learning. In particular, lexical guessing is taken as indispensable to compensate for the limited vocabulary size of second/foreign learners and thus greatly promoted for reading comprehension (Twaddle, 1973). Since it is unrealistic to expect learners to acquire a large vocabulary in a second/foreign language, lexical guessing is considered a supplementary skill to prepare them to cope with such tasks as reading and listening. When encountering unknown words or unfamiliar words in a context, learners can be trained to infer the meaning of a word based on the clues provided by the context. This technique is especially useful for dealing with low-frequency words. Therefore,

by definition, lexical guessing means to make a guess, a "sensible guess", at the meaning of an unknown lexical item based on contextual clues, which may include the syntactic structure of the sentence in which an unknown item occurs, or real-world knowledge contained in the text (Twaddle, 1973: 72).

Before any further discussion, it is important to note that a few terms are used interchangeably in the literature and refer to the same concept. These include contextual guessing (van Parreren & Schouten-van Parreren, 1982), guessing from context (Haynes, 1993: 50), inferring from context (Sokmen, 1997: 237), inferencing from context (Schmitt, 2000: 152), lexical inferencing (Haastrup, 1991: 39) or inferencing (Fraser, 1999: 230; Paribakht & Wesche, 1999: 214). Although the term "contextual guessing" has been preferred in the previous sections of this chapter, "lexical guessing" or "lexical inferencing" will be used in this section in order to emphasize that it is the meaning of a lexical item, rather than of a text which is the focus of the learner's effort. Used alone, the term "inferencing" has a broader sense referring to the strategy commonly used to construct linguistic hypotheses in language learning (Bialystok, 1983: 105), or is taken as all the strategies and skills that a reader adopts in an attempt to make sense of a text (Kembo, 2001: 77). Therefore, it would be more appropriate to use the term "lexical inferencing", rather than "inferencing" as used by some other applied linguists.

To define lexical guessing, it is the process involving guessing the meaning of a word on the basis of "all available linguistic cues in combination with the learner's general knowledge of the world, her awareness of the co-text and her relevant linguistic knowledge" (Haastrup, 1991: 40). Based on this definition, knowledge sources can be further discriminated and both the benefits and drawbacks associated with them will be discussed by presenting recent studies and research articles.

Three types of clues in relation to lexical guessing are distinguished, including intra-lingual, inter-lingual and contextual clues (Carton, 1971: 50; Haastrup, 1991: 92). Intra-lingual clues refer to knowledge of the target language, and may include knowledge about word classes, and morphological and syntactical structures. Inter-lingual clues are information about the native language or other foreign language other than the target language. Contextual clues are either non-linguistic, extra-linguistic clues which refer to the world knowledge we have or the information provided by the linguistic context in which a target word is embedded. These three types of clues may act individually or interactively and contribute to lexical guessing in particular and to the comprehension process in general. However, studies show that not all the three types of clues are necessarily conducive to lexical guessing.

A few studies find that word-level analysis, based on either intra-lingual or inter-lingual clues, will frequently lead to misunderstanding. That is, problems are likely to occur when the learner bases lexical guessing on morphological clues rather than on contextual clues (Bensoussan & Laufer, 1984; Haynes, 1993). Intra-lingually, words that have similar lexical forms or morphological structures to target words can also cause problems. These are words which learners think they know but actually misunderstand on the basis of inadequate morphological analysis or because they mistake a similar word for a target word. The former type of word is called a "deceptive transparency" by Laufer (1997: 25), and the latter a "synform" by Laufer (1997: 26) or a "mistaken ID" by Huckin & Bloch (1993: 160). An example of a deceptive transparency might be "infallible", which is seen as composed of three morphemes, *in-*, *fall* and *–ible* (Laufer, 1997: 25) and mistakenly understood as "something that cannot fall"; and "cute", when understood as "acute", would be an

example of a synform (1997: 26). Therefore, learners may be led to misunderstand the message conveyed, and be unable to construct an appropriate meaning.

Inter-lingually, the existence of cognates, for example, is a major problem that is frequently discussed in the literature. Cognates can help learners comprehend reading texts if the meaning of a cognate is related to the meaning of a target word; however, they may impair understanding if they are not semantically related to the L2 words (Haynes, 1993; Meara, 1993; Laufer; 1997). As a result, reliance on cognates cannot be taken as a practical strategy both because some cognates, so called "false friends" are misleading, (Bensoussan & Laufer, 1984; Meara, 1993; Laufer, 1997; Swain, 1997), and because they only exist in languages within the same language family and are not particularly useful for unrelated languages (Koda, 1997: 39).

In previous sections, the discussion of the role of context in vocabulary learning showed that lexical guessing based on contextual clues is a major strategy for reading comprehension. These clues can be located in the immediate context, or beyond the immediate context. The immediate context is what Haynes calls the "local context", referring to the sentence where the target word appears, while the "global context" refers to the sentences or clauses beyond the sentence containing the target word (Haynes, 1993: 52). Clues in the immediate context may be easier for the learner to locate, whilst those beyond the immediate context require abilities such as relating the logic of the passage as a whole, or abstracting the main idea of the co-text. The latter is certainly not easy for many learners and thus requires training.

Without instruction or training, wild guessing is very likely to occur and poses risks for vocabulary development because wrongly guessed word meanings may be

retained in learners' memories and become difficult to eliminate. The empirical studies cited in section 2.2.3 demonstrate this kind of danger (Dubin & Olshtain, 1993; Haynes, 1993; Huckin & Bloch, 1993; Meara, 1993; Laufer, 1997). As argued, lexical guessing may cause lemmas and lexemes to be incorrectly matched in the mental lexicon. A lemma is a conceptual unit, which specifies the semantic and syntactic properties of a lexical item whilst lexemes represent the phonological and morphological features (de Bot, Paribakht & Wesche, 1997: 313). Therefore, a wrongly guessed word will assign an inappropriate lemma to the lexeme encountered by the learner.

To avoid the danger of wild guessing or incorrect guessing, learners need to be taught how to make use of various contextual clues in order to compensate for the shortcomings of lexical guessing (Huckin & Coady, 1999). The learner can be trained in lexical guessing (van Parreren & Schouten-van Parreren, 1982). Procedures are suggested by Nation and Coady (1988) which include 1) analysis of the part of speech of the unknown word, 2) looking for any clues near to the word, 3) looking for any clues embedded within a wider context such as clauses and sentences, 4) guessing its meaning, and 5) double-checking the meaning to see whether it fits into the text (Nation & Coady, 1988: 104-5).

Nonetheless, the success of lexical guessing involves other factors. The problem of word guessability is highlighted in Na & Nation (1985). Their investigation revealed that linguistic factors such as word density and part of speech would influence the success of lexical guessing. Word density refers to the ratio of unknown words to known words. With the same number of unknown words, a long passage was found to be more understandable than a short passage of the same content because of its low density. With respect to parts of speech, verbs and nouns

were found to be easier to guess than adjectives and adverbs. In order to attain a

higher success rate in lexical guessing, they further suggested that the activity of

lexical guessing be conducted as group work in the initial stage rather than an

individual exercise to enable the learners to become familiarized with the strategy

(Na & Nation, 1985: 39).

An additional problem affecting word guessibility was noticed by Bensoussan

and Laufer (1984). "Preconceived notions" referred to inappropriate meanings about

an unknown or unfamiliar word, which were frequently imposed on the word by the

learner in understanding a text (Bensoussan & Laufer, 1984: 22). This assumed

meaning would prevent the learner from making use of available clues.

Consequently, the learner would be liable to misunderstand the meaning conveyed in

the text.

Laufer summed up four situations that will affect the effectiveness of lexical

guessing (1997). First, it is naïve to expect that the context will always provide the

necessary clues for lexical guessing; indeed, sometimes contextual clues do not exist.

Second, additional unknown or unfamiliar words other than the target words are

likely to hinder the process of lexical guessing because they are likely to interrupt the

process of the reconstruction of the message. Third, as discussed, the presence of

distracting clues such as the existence of cognates may be misleading in lexical

guessing. Fourth, learners are liable to impose the interpretation derived from their

world knowledge on the information actually transmitted in the message whenever

there is a mismatch between what they already know and the message the text

conveys (Laufer, 1997: 31).

In short, despite the problematic issues highlighted above, lexical guessing is a

major strategy for the learner to deal with unknown words in a text, and especially

with words of low frequency. In particular, when contextual clues are used properly, it is regarded as a more proper and constructive way to achieve reading comprehension or even build vocabulary knowledge because too much reliance on current vocabulary knowledge is not so reliable and may even be detrimental. No matter which clue is used, encouraging the learner to check guesses against context after the process of lexical guessing is necessary (Nation & Coady, 1988; Haynes, 1993; Huckin & Bloch, 1993; Hulstijn, 1993).

2.3.2 Dictionary Use

The second reading strategy, dictionary use, is generally believed to interrupt reading comprehension because it requires a substantial amount of time both in consulting the dictionary and in choosing an appropriate lexical entry for polysemous items where multiple meanings are given. Before investigating the relationship between dictionary use and the acquisition of word meaning, a distinction between dictionary use for comprehension and for vocabulary learning needs to be made.

Dictionary use for comprehension is often discussed in studies of second language reading, but is always discouraged by language teachers because of the consumption of reading time, as stated above. Although it has been found that dictionary use does affect reading speed, this does not automatically imply that it also impedes reading comprehension (Knight, 1994; Fraser, 1998). Contrary to general belief, access to a dictionary could increase reading comprehension. In his study, Knight found that dictionary use increased comprehension and enhanced vocabulary learning. Perhaps unsurprisingly, it was the learners of low L1 verbal ability (as measured by the American College Test) who benefited most from access

to dictionaries (Knight, 1994). They performed significantly better both in comprehension and vocabulary tests than no dictionary groups with the same verbal ability.

Similarly, Fraser in her review of empirical studies concluded that consulting a dictionary during the process of reading was found not only to enhance reading comprehension but also contributed to vocabulary development (1998). Three functions of this strategy were considered to be crucial for long-term memory. First, the learner's attention is focused on the form-meaning connection while looking up a word in a dictionary. Noticing this relationship is the first step to learning a new word. Second, while searching for the meaning of a word, the learner actually rehearses the word and thus forms a temporary representation in working memory. Rehearsal offers an opportunity for information to stay in the memory longer. Finally, after locating the word in a dictionary, the learner has to choose the proper part of speech to fit into the text, make an analysis of its grammatical properties, and even relate it to existing knowledge or to L1 words. This is an elaboration which involves deep processing and creates deeper traces in memory (Fraser, 1998: 84). These three functions provide a better chance for an unknown word to be retained in long-term memory.

Additionally, the issue of the quality or usefulness of different types of dictionaries has been the subject of a vigorous debate, especially between the use of monolingual vs. bilingual dictionaries in relation to second language learning. Monolingual dictionaries are generally considered superior to bilingual dictionaries, as pointed out by Bejoint (1987: 99). The inadequacy of bilingual dictionaries in establishing the learner's second/foreign language lexicon lies in the impossibility of finding absolute equivalents in different languages (Snell-Hornby, 1984: 275). The

translation equivalents of the native language used in bilingual dictionaries approximate to the meaning of the target language lexical item only to a degree and this problem is worsened when taking into account the factor of the contextual use and pragmatic functions of the lexical item (Snell-Hornby, 1984: 276-8). The degrees of approximation depend on different domains, with a higher degree existing in lexical items referring to common terms and the lowest degree in those containing cultural-related elements (Snell-Hornby, 1987: 165-6). In addition, highly polysemous words require context if the reader is to recover the intended meaning (Bejoint, 1987: 101). These arguments cast doubt on the effectiveness of bilingual dictionaries in that the effort of trying to pin down a definite meaning for a foreign word with translation equivalents is neither logical nor practical.

A further objection provided out by Baxter is that, if the bilingual dictionary provides equivalents in the native language, the learner does not have the opportunity to develop the ability to define a word (1980). A major benefit of using monolingual dictionaries is that it can train the learner to develop the ability to define words by reading the definitions provided in monolingual dictionaries. It is therefore claimed that the capability of monolingual dictionaries to promote this ability surpasses the usefulness of bilingual dictionaries (Baxter, 1980: 330).

However, the preference for bilingual dictionaries over monolingual is found to be so robust (Baxter, 1980; Hartmann, 1987; Laufer & Kimmel, 1997) that it is not only difficult but also often impractical to force the learner to give up the habit of using bilingual dictionaries (Hartmann, 1987: 22). It is not only learners of low and intermediate proficiency who rely heavily on bilingual dictionaries, but also advanced learners, who consult this type of dictionary more often than their monolingual counterparts (Tomaszczyk, 1983: 46). It is assumed that affective

factors are involved in this kind of strong preference for bilingual dictionaries in that learners feel more secure when accessing translations or explanations in their native language (Laufer & Kimmel, 1997: 362). This may be because one major obstacle in the use of monolingual dictionaries is that the learner cannot easily understand the definitions or target language synonyms. The definitions are especially difficult for learners of low and intermediate proficiency (Bejoint, 1981; Horsfall, 1997; Laufer & Kimmel, 1997; Gonzalez, 1999). To solve this difficulty, the idea of using a "bilingualised" dictionary is proposed. A "bilingualised" dictionary displays both target language definitions and native language translation equivalences simultaneously (Laufer, & Kimmel, 1997: 363), in contrast with bilingual dictionaries which do not contain target language definitions. By providing both kinds of information, this type of dictionary is argued to be able to meet the needs of most language learners, whether they choose to focus on or read the bilingual or the monolingual part.

Dictionary use for vocabulary learning is an intriguing aspect of our discussion of vocabulary development. First of all, the value of dictionary use in developing independent vocabulary learning is worth mentioning (Horsfall, 1997; Gonzales, 1999). In interviews with second language learners, Gonzales found that dictionary use is a frequent strategy pursued by learners when encountering unknown words in a text and therefore cannot be neglected since it satisfies the learners' vocabulary needs in respect of both breadth and depth knowledge (Gonzales, 1999: 269).

Moreover, Luppescu and Day revealed that dictionary use has a facilitating effect on short-term word retention (1993). In their study, subjects in treatment groups were allowed to use dictionaries while reading a short story, whereas those in the control group were not. An immediate vocabulary test was conducted and the

result showed that the treatment group scored significantly higher than the control group. However, they admitted that further investigation was needed to test the long-term retention of word meaning as a consequence of dictionary use (Luppescu & Day, 1993: 277).

Empirical research regarding the effect of dictionary use conducted by Hulstijn *et al.* (1996) reveals further important information. Three groups of learners were subjected to three different learning conditions: provision of marginal glosses, access to dictionaries and a control group without any aid while reading (see section 2.2.3). Although it was found that the performance of learners with access to marginal glosses was significantly better than that of the other two groups, the scores for word retention for those words which were looked up by learners with access to dictionaries were actually higher than those of learners in the marginal gloss group. In particular, two important findings are worth noting. First, learners rarely looked up unknown words even when they were allowed to use a dictionary. Only when they perceived the relevance of an unknown word to overall understanding of a text did they make an attempt to look up the words. Second, this research proved that dictionary use could enhance vocabulary learning in that it engaged the learner in a form-meaning relationship which is considered the initial step of vocabulary development (Hulstijn *et al.*, 1996).

The extent of the learner's needs or motivation to use a dictionary may provide vital information about how and when to promote such a strategy. This is a major concern of Hulstijn in another study (1993). By adopting a computer-based procedure, the researcher simultaneously investigated two types of variables, task variables and learner variables, and questioned the extent to which they affect learners' look-up behaviour. As the results show, two task variables, i.e. the

perceived word relevance to reach reading goals and the task type, and the existing vocabulary knowledge of the learners were confirmed to have significant effects on learners' behaviour. If we are to promote dictionary use in reading, these factors need to be taken into account.

To help the learner use a dictionary effectively, a closer examination of the look-up process may enlighten us with regard to the process of reading. As revealed by Schofield, several steps are involved in the look-up process during reading. Apart from the initial stages of looking up target words following the alphabetical order of the dictionary, an inference is also required to differentiate among the polysemous senses of most words and choose an appropriate one to match the context (Schofield, 1982: 189). In other words, teaching the learner about how to use a dictionary effectively is a crucial step before recommending dictionary use. Moreover, the strategy requires much practice. Furthermore, the learner can be advised to guess the meaning of an unknown word before using a dictionary, then verify this inferred meaning by consulting a dictionary, and then repeat the word and elaborate its meaning to create a deeper trace in memory. Empirical studies have proved that higher retention of word meaning occurs when the learner first attempts to infer the meaning of an unknown word before consulting a dictionary (Fraser, 1999).

In brief, most researchers recognize a place for dictionary use in reading comprehension and vocabulary learning, be it a monolingual or a bilingual dictionary. As we have seen, learning words from context can produce greater effects on vocabulary learning when combined with the use of dictionaries because a dictionary provides a valuable opportunity to verify the word meanings inferred from the context. Thus, verification is widely regarded as a crucial step in vocabulary development (Nation, 1990; Mondria & Wit-de Boer, 1991; Hulstijn, 1993). Finally,

knowing how to use a dictionary is not sufficient for vocabulary development. As Horsfall concluded, it is of more importance to instruct learners in ways of using a dictionary that enable them to transfer knowledge from the dictionary to learners' memory and store it in long-term memory for subsequent use (Horsfall, 1997: 8)

2.3.3 Ignoring

In the model proposed by Gass (1988) (see chapter 1), the first step in the acquisition process is apperception. For vocabulary learning, any lexical items that are ignored or fail to be noticed by language learners cannot therefore be processed as input or integrated into the learners' existing knowledge structure. As is apparent from many papers, this strategy is commonly adopted by learners (Bensoussan & Laufer, 1984; Dubin & Olshtain, 1993; Haynes, 1993; Huckin & Bloch, 1993; Hulstijn *et al.*, 1996; Paribakht & Wesche, 1997, 1999; Huckin & Coady, 1999). For example, the subjects in Paribakht & Wesche's study ignored an unexpectedly high percentage of over 50 % of unknown words (Paribakht & Wesche, 1999). Similarly, the Dictionary group in Hulstijn *et al.* were found to ignore some unknown words and not to look up their meanings even though they were allowed to use dictionaries (Hulstijn *et al.*, 1996). Both these and other comparable studies account for such a failure to notice in terms of the learners' perception that these ignored words are not relevant to the task demand. For effective vocabulary learning, classroom practice needs to direct learners' attention to the target words (Paribakht & Wesche, 1997).

Viewed from a different point of view, ignoring may be both useful and practical, a point made in several studies (Nation, 1990; Oxford & Scarcella, 1994; Fraser, 1999). Nation, for example, suggests that for low-frequency words, ignoring can be practical (Nation, 1990: 133), for it is time-consuming and interruptive to

look up every unknown word. Learners should therefore be taught when to ignore unknown words. This is because reading is a selective process (Goodman, 1976: 260), during which learners learn how to choose information relevant for comprehending the passage based on their expectations.

To conclude this section on reading strategies, it therefore seems worthwhile to train learners how and when to adopt useful and practical reading strategies to deal with unfamiliar or unknown words (Fraser, 1999). Fraser reported a study which explored the effect of *Lexical Processing Strategies* (LPS) instruction (1999). Three strategies were involved: ignoring, consulting (dictionaries or experts), and inferring (1999: 226). These three strategies were first introduced to the learner so as to establish what they were, how they worked and why they were important in dealing with unknown words. This was followed by a phase of practice in using these strategies. Think-aloud protocol data were collected by interviewing individual learners. The result of the data analysis revealed important information with respect to the frequency of use of each strategy in reading, the respective effectiveness of the strategies, and their association with reading comprehension and vocabulary learning. It showed that reading comprehension increased as a result of successful determination of word meanings using the strategies of consulting and lexical inferencing. Moreover, when these two strategies were used together, the retention rate was found to be higher than when using single strategies. It was therefore recommended by the researcher that instruction on the use of these three strategies enabled the learner to adopt a suitable strategy or combination of strategies for dealing with unknown words and thus indirectly enhanced vocabulary learning (Fraser, 1999: 236).

2.4 Aspects of Vocabulary Knowledge

2.4.1 Knowing a Word

As early as 1976, Richards reminded us of various properties of vocabulary knowledge that needed to be taken into account in assessing vocabulary knowledge. Specifically, knowing a word included understanding 1) its frequency and collocation[2]; 2) the stylistic constraints on the use of registers; 3) the structural and grammatical properties of the word; 4) base forms and morphological rules contributing its derivative and inflected forms; 5) words associated in its semantic network; 6) semantic features and corresponding values; and 7) polysemous senses.

Nation classified vocabulary knowledge into four general categories: form (spoken and written), position (grammatical patterns and collocation), function (frequency and appropriateness) and meaning (concept and association) (Nation, 1990: 31). In total, eight subcategories were identified, with receptive and productive knowledge being further distinguished in each. For instance, receptive knowledge of spoken forms involves the ability to make a distinction between phonetically similar words, and productive knowledge involves the ability to produce the sounds.

Though they each provide different taxonomies, both Richards and Nation call for more attention to the fact that knowing a word is not merely a matter of knowing its meaning, and accordingly we should not restrict our learning and teaching of L2 vocabulary to meaning alone but instead should also pay attention to relevant linguistic knowledge.

---

[2] Richards mentioned these two properties in his assumption 1 as though they were related (1976: 79).

49

Following the distinction made by Bialystok and Sharwood-Smith (1985), vocabulary learning will be discussed in terms of knowledge and control. This distinction will enable us to categorize theories of vocabulary knowledge.

2.4.2 Knowledge: breadth and depth

Generally speaking, researchers agree on the important distinction between vocabulary breadth and depth. Vocabulary breadth refers to the size of the vocabulary or the total numbers of words known to language learners, while vocabulary depth refers to the quality of vocabulary knowledge possessed by learners (Nation, 1990; Wesche & Paribakht, 1996). Size is the quantative aspect and depth is the qualitative aspect. Vocabulary learning research typically focuses on either breadth or depth, with very few studies measuring both as indication of vocabulary knowledge.

A more recent paper by Henriksen classifies vocabulary knowledge into three dimensions: 1) partial to precise knowledge; 2) depth of vocabulary knowledge; and 3) receptive and productive knowledge (Henriksen, 1999: 304). The first dimension, partial to precise knowledge, ranges from word recognition to accurate understanding of lexical items. It is presented as a continuum with one extreme representing word recognition and the other precise comprehension. Many levels may be involved, indicating different degrees of word knowledge. According to Henriksen, this dimension is seen as a mapping process, i.e. mapping meaning onto form. "Item learning" characterizes this process during which vocabulary size is being expanded (1999: 310). The second dimension, depth of vocabulary knowledge, is mainly related to word association. Knowing a word entails knowing its various types of relations to other words, including paradigmatic and syntagmatic

associations. All possible associated words are connected to each other in a semantic network and therefore the whole semantic system is under constant restructuring. Dimensions 1 and 2 constitute the process of "semantization", which accounts for the process of increasing vocabulary knowledge (1999: 307-8). The final dimension is related to receptive and productive knowledge. Different from the previous two dimensions, this dimension basically concerns the ease with which learners access or use words, and is therefore considered to be linked to the control aspect of vocabulary knowledge.

Though theoretically distinguished, vocabulary breadth and depth may be related to each other in some way. Two studies have set out to explore their relationship. One study was conducted by Schmitt & Meara (1997), who tested subjects on vocabulary size, word association, suffix knowledge and general language proficiency. For the vocabulary size test, they selected two established tests as measuring instruments: the Eurocentres Vocabulary Size Test (Meara & Jones, 1990) and the Vocabulary Levels Test (Nation, 1990). To measure vocabulary depth, they devised a method to test word associations and suffix knowledge, both having receptive and productive versions. By using a carefully selected word as a stimulus, they tried to elicit its word associates, as well as derivative and inflected forms. General language proficiency on the other hand was assessed with the TOEFL Test. The result revealed that word associate knowledge correlated with vocabulary size ($r=0.62$ for the productive version, and $r=0.61$ for the receptive version of the test). The relationship between suffix knowledge and vocabulary size was weaker, however. It is clear therefore that at least one aspect of vocabulary depth is closely related to vocabulary size.

Another piece of empirical evidence can be found in Nurweni and Read's study (1999). They used different tests to measure learners' vocabulary knowledge: a translation task for vocabulary size, a word associates test for vocabulary depth (Read, 1993) and an interview to explore deeper word knowledge. The major finding was that the score for word associates modestly correlated with the score for the translation task with $r = 0.62$. They interpreted this as indicating that vocabulary depth and breadth are actually related.

Though using different tests to measure the correlation between vocabulary breadth and depth, these two studies arrive at the same conclusion. However, one may doubt the reliability of these two tests so that more research is required to present a clearer picture of the suggested link of vocabulary breadth and depth.

2.4.3 Control: receptive and productive knowledge

As has been discussed, Nation made a distinction between receptive and productive knowledge, and considered the latter a more advanced kind of knowledge than the former (Nation, 1990: 31-2). That is, productive knowledge exists only when the very first stages of learning have been completed and continues to lag behind receptive knowledge thereafter. Receptive vocabulary knowledge is also referred to as passive knowledge, while productive knowledge is considered as active knowledge, which implies that this kind of knowledge can be accessed automatically. In fact, there is little disagreement on this point.

To distinguish receptive/passive from productive/active vocabulary knowledge, Melka postulates that receptive and productive knowledge can be seen as two poles in a continuum of vocabulary knowledge (Melka, 1997). The notion of familiarity with words is conceived as the basis of this distinction (1997: 85). With increasing

language proficiency, the learner's ability will move toward the pole of productive knowledge. That is to say, a higher degree of vocabulary knowledge is evidenced in the form of productive knowledge, or can be more readily transformed into productive knowledge. However, for Meara, the distinction between receptive and productive knowledge is a qualitative rather than a gradual one. He argues that a passive vocabulary item is conceived of as an item in a semantic network with unidirectional associations such that mention of other items in the network prompts association with the passive item, but mention of the passive item does not prompt association with other items (Meara, 1990). Therefore, prompting learners with the passive item alone is unlikely to connect it to an active item. Rather, "exercises which deliberately stress the associational links leading from already known [i.e. active] words to newly learned [i.e. passive] words might be a more effective way of activating a passive vocabulary" (1990: 153).

Despite these few notable attempts to distinguish receptive and productive knowledge, most research under review in this paper does not attempt this, or is not able to make a clear boundary between receptive and productive knowledge. Some studies test word recognition after experimental treatment; some test both recognition and recall of new vocabulary but without providing any explanation of what theoretical basis underlies these two kinds of measure in relation to receptive and productive knowledge.

As with vocabulary breadth and depth, a similar question comes to mind with respect to receptive/passive and productive/active knowledge, namely whether they are related. A study by Laufer & Paribakht (1998) looked into this question, first by defining three different kinds of word knowledge: passive, controlled active and free active. Passive knowledge referred to the knowledge of a word's most common

meaning; controlled active was a type of productive knowledge elicited by cues; and free active knowledge was the ability to use words automatically in free writing (Laufer & Paribakht, 1998: 370). Again, the Vocabulary Levels Test was used to measure vocabulary size. To test controlled active vocabulary, the Productive Vocabulary Levels Test was used, where a target word in a sentential context was cued by supplying its first few letters (Laufer & Nation, 1999). The Lexical Frequency Profile (Laufer & Nation, 1995) was used to determine the lexical richness in a composition of 300 to 400 words by calculating the percentages of words from various frequency levels (1000, 2000, 3000, 5000 University Word List and 10000). They further broke down the scores for different frequency levels and found that the gap between passive and controlled active knowledge was smaller for words at higher frequency levels. Put another way, learners could both recognize and use more common words in a controlled manner. Additionally, significant correlations between controlled active and passive, free active and passive, and free active and controlled active were also detected. For the EFL group (Israeli students), controlled active and passive gave a 0.89 ($p<$ .001) coefficient; free active and passive, $r=0.68$ ($p<$ .001); free active and controlled active, $r=0.68$ ($p<$ .001). For ESL (Canadian students), the results for the association of controlled active and passive were 0.72 ($p<$ .01), although free active and passive, and free active and controlled active had lower correlations at $r=0.41$ ($p<$ .01) and $r=0.51$ ($p<$ .01) respectively (Laufer & Paribakht, 1998: 378). Moreover, a comparison between different proficiency levels (low, intermediate and advanced based on the scores of passive knowledge) was made to see whether the development of passive knowledge was accompanied by increased controlled active and free active knowledge. The results confirmed this speculation.

For the time being, given the positive evidence found, it seems reasonable to conclude that receptive/passive and productive/active knowledge are interrelated. Although more empirical research will be required to clarify the nature of this relation, an issue not directly addressed in the present study.

2.4.4 Vocabulary Knowledge and Reading Comprehension

In L1 research, it has been claimed that level of vocabulary knowledge is strongly associated with level of reading comprehension (Nagy & Herman, 1987; Sternberg, 1987). Nagy & Herman emphasize that multiple exposures result in a large quantity of vocabulary growth and depth, which in turn facilitates subsequent reading comprehension (Nagy & Herman, 1987: 31). In foreign language reading, the "threshold hypothesis" claims that a certain amount of vocabulary knowledge is needed for minimum comprehension (Alderson, 1984: 18). For learners of low proficiency, it is particularly true that the L2 reading problem always becomes a language problem. Therefore, we speculate that there is a close relationship between vocabulary knowledge and reading comprehension. That is, the extent of vocabulary size and depth can predict the level of reading comprehension.

Nevertheless, in L2 research, very few studies have been dedicated to investigating this relationship. Among them, Laufer specifically researched the relationship between vocabulary size and reading comprehension (Laufer, 1992). She used Nation's Vocabulary Levels Test (Nation, 1990) to test one group of subjects, and Meara & Jones' Eurocentres Vocabulary Test (1990) to measure the vocabulary size of the other group of native Arabic speaking university students. The first group also took a reading test taken from *Examen Hoger Algemeen Vortgezet Onderwijs* and the second took the comprehension sub-test taken from the

Israeli university psychometric entrance test. High correlations were found between vocabulary size measures and reading comprehension tests. It was thus concluded that vocabulary size could predict reading performance. Furthermore, Laufer also made an attempt to associate scores on the vocabulary level test with reading ability, and found that subjects reaching the vocabulary level of 3,000 would be likely to achieve successful reading comprehension. This means that the 3000 level was the threshold of effective comprehension (Laufer, 1992: 129). Below this level, reading would tend to be difficult. In addition, the difference of reading scores at different levels was found to be significant. This implied that different levels of vocabulary knowledge could be a decisive factor for reading comprehension.

A recent study by Qian was deliberately set up to investigate the possible interrelationship between vocabulary size and depth, and reading comprehension (Qian, 1999). Measuring instruments included the Vocabulary Levels Test (Nation, 1990) for vocabulary size; the Word Associate Format devised by Read (1993) for assessing depth of vocabulary knowledge in terms of meanings and collocations of target words, and words paradigmatically and syntagmatically related to these words; and a reading comprehension test adopted from TOEFL, but with only four reading passages and corresponding comprehension questions. In addition, Qian measured morphological knowledge with a Morphological Knowledge Test devised by himself. This measured the ability of the learner to define a target word and its affixes, and the ability to identify the part-of-speech of the stem word and its affixes (Qian, 1999: 293). The result showed that at a 0.05 level of significance, the correlation between vocabulary size and reading comprehension was 0.78; between depth of vocabulary knowledge and reading comprehension the correlation was 0.82; and between morphological knowledge and reading comprehension the correlation was 0.64.

Accordingly, it was confirmed that vocabulary knowledge, both breadth and depth, was significantly associated with reading comprehension. In other words, a similar conclusion could be drawn, that in L2 reading, quantity and quality of vocabulary knowledge can predict comprehension.

Crucial findings were obtained in a study by Schoonen, Hulstijn and Bossers (1998) with respect to the effect of language-specific knowledge (i.e. vocabulary knowledge of native and foreign languages) on native and foreign language reading comprehension among Dutch students, (although the research goal of the experiment was to identify the level at which the instruction in metacognitive knowledge and strategies, i.e. general reading skill, would be most useful). Subjects included grade 6, 8, and 10 students. The researchers used a standardized multiple-choice test containing narratives and expository texts but gave the subjects different materials based on their grades in a foreign language reading test. As for testing vocabulary knowledge in a foreign language, i.e. English, a translation test was given, which was composed of English sentences with Dutch translations provided but leaving blanks for target words. The subjects were supposed to translate the words into Dutch. Metacognitive knowledge, on the other hand, was measured by questionnaires consisting of four elements: assessment of oneself as a reader, knowledge of reading goals and comprehension criteria, knowledge of text characteristics and knowledge of reading strategies, each containing a number of questions (Schoonen *et al.*, 1998: 79). The result found slight differences among different levels. For students in grade 8 with three years of foreign language instruction, vocabulary knowledge of English was found to predict the level of reading comprehension ($r=$ 0.72), whereas for grade 10 students who had received five years of instruction, vocabulary knowledge of English also affected reading

comprehension but not to the same degree as for grade 8 students ($r=0.38$). In other words, for grade 10 students, the importance of metacognitive knowledge of reading strategies increased while the importance of vocabulary knowledge decreased. This implies that a certain level of vocabulary knowledge of a foreign language is essential for reading comprehension. But beyond this level, the importance of vocabulary knowledge diminishes in relation to general reading ability. The transfer of L1 reading ability thus becomes possible at this stage (Schoonen *et al.*, 1998: 101).

To sum up, the available evidence shows that reading comprehension is influenced by vocabulary knowledge. Researchers have repeatedly found that the level of vocabulary knowledge can be used to predict performance on reading comprehension.

2.5 Conclusion

In this chapter, we have reviewed a considerable number of research studies regarding second language vocabulary learning. These include research into vocabulary learning from context by means of reading (i.e. incidental vocabulary learning), reading strategies that can contribute to vocabulary learning, and the components of vocabulary knowledge that are generally agreed as crucial aspects in second language vocabulary development. A number of important issues are involved in this review. First, although context is seen as helpful for vocabulary learning, problems exist with respect to the provision of context clues, which are not always transparent and usable. It is suggested that teachers should be cautious in encouraging learning from context. Second, reading strategies that are used to manage unknown or unfamiliar words are worth noting because a combination of these strategies is regarded as conducive for vocabulary development. Third, it is

important to understand the dimensions of vocabulary knowledge, including breadth and depth, and receptive and productive knowledge. In other words, the manifestation of vocabulary knowledge is not as simple as has been thought. Finally, the research literature shows that vocabulary knowledge and reading comprehension are closely related. One of the major purposes of this literature review is to draw attention this close relationship. The discussion of these topics will be summarised in more detail in the following sections.

2.5.1 The Contextual Factor

From the discussion in section 2.2, we know that context is influential not only for reading comprehension, but also for vocabulary learning. By using available contextual clues, learners can successfully achieve comprehension even when some unknown or unfamiliar words are embedded in the text. More importantly, context provides a mental environment for learners to be engaged in some kind of mental work so as to lay down a deeper trace in the memory system and thus to enhance vocabulary retention (Craik & Lockhart, 1972; Goodman, 1976). Additionally, the opportunity provided by reading for multiple exposures to words makes this approach highly recommended for vocabulary learning (Nagy *et al.*, 1985; Krashen, 1989).

On the other hand, several kinds of danger in using contextual clues have also been discussed, including "non-existent", "unusable" or even "misleading" contextual clues (Haynes, 1993; Laufer, 1997). It is not that the learner does not notice or utilise these clues, but rather that he/she is not able to detect their usefulness, or may even misuse them. This is because the learner's existing knowledge is insufficient. Therefore, a theory of threshold vocabulary knowledge

for reading is proposed to alleviate the problem and a careful consideration of appropriate reading materials is required.

Ways of modification of the text presented to the learner have also been considered. As some empirical studies show, the provision of marginal glosses has proved to be a promising method for modifying context; in particular, glosses in a multiple-choice format are shown to be valuable for vocabulary learning (Davis, 1989; Jacobs *et al.*, 1994; Hulstijn *et al.*, 1996)

2.5.2 Reading Strategies and Vocabulary Learning: the learner factor

A second factor in the reading process is the L2 learner, and especially the strategies they adopt to deal with unknown or unfamiliar words. The three major reading strategies most frequently used by the learner to cope with unknown words during reading have been discussed. First, the effect of contextual clues discussed in the previous section is also relevant to the reading strategies adopted by learners. Contextual guessing is a strategy that engages the learner in guessing the meaning of an unknown word based on contextual clues. Though in general it is considered influential and is greatly promoted for reading, this strategy seems less suitable for low-proficiency learners. Even when learners are able to learn from context, the ability to use different kinds of contextual clues needs to be taught to avoid wild guessing and hence inappropriate retention. Second, dictionary use is one way to provide the opportunity for verification of guessed words, but should be encouraged for selected words, rather than for every unknown word because it interrupts the reading process. Finally, the strategy of ignoring should only be used for low-frequency words in the text (assuming that learners can identify unknown items as low frequency) and words that are not perceived to be important for understanding a

text. These strategies should not be applied randomly but in a principled way. Skills need to be taught and awareness should be raised as to when to choose an appropriate strategy in a particular situation. This should be the main point when training students in reading strategies.

2.5.3 Aspects of Vocabulary Knowledge: the word factor

Last but not least, the word factor can influence reading comprehension to a great extent. In a broader sense of word knowledge, it includes both breadth and depth of vocabulary knowledge. As has been discussed, each aspect will affect reading comprehension to some extent. Most studies agree that a larger size of vocabulary and a deeper knowledge of words are conducive to successful reading comprehension since aspects of vocabulary knowledge are manifest in the learner's proficiency level. Therefore, the acquisition of vocabulary knowledge should not be limited to one aspect of word knowledge, but needs to cover all related dimensions.

As a matter of fact, these three factors do not function alone; they are closely related to each other. For example, the proficiency level of the learner can affect the understanding of the contextual clues needed to guess the intended meaning of an unknown lexical item. Moreover, the perceived relevance of an unknown word to the overall understanding of the text can possibly determine which strategies, including guessing, dictionary use, or ignoring, will be adopted by the learner while reading a text.

2.5.4 The Place of the Current Research

The review of the literature concerning vocabulary learning presented in this chapter suggests that the precise nature of the interrelation between reading and

vocabulary learning has not been demonstrated empirically. Despite the claim that reading is important for incidental vocabulary learning, it has not been illustrated in what precise way this occurs. That is, the specific mechanism that contributes to vocabulary learning in reading has not been investigated in any of these studies. Moreover, although it has been proved that both the quantity and quality of vocabulary knowledge are associated with the level of reading ability, the existing studies have not uncovered precisely how reading comprehension is associated with vocabulary knowledge. In other words, while we believe that vocabulary knowledge and reading comprehension are interdependent and have been rightly regarded as two sides of a coin in second language learning, it has not been demonstrated empirically that reading helps the learner acquire vocabulary knowledge, and especially vocabulary depth knowledge. The lack of such an investigation provides a place for the current study, which is intended to justify the intuition that reading and vocabulary learning are interdependent and, in particular, to reveal specifically that reading a text helps establish vocabulary depth by appealing to the way in which the lexical items occurring in a text are related. To acquire vocabulary depth requires understanding these relations as they appear in a text.

In addition, it is well established that context is helpful for lexical guessing by providing various kinds of clues as to the likely meanings of unknown items. However, there has never been a principled definition of the notion of context. That is, context remains an abstract notion without detailed specification and lacks an operational definition that can be tested empirically. Although investigated in earlier research, the notion of "contextual richness" is obscure and the criteria provided by existing research are disparate (e.g. Mondria & Wit-de Boer, 1991; Zahar *et al.* 2001). In addition, nothing has been mentioned about the structure of contexts. This

constitutes a starting point for the current research. In this research, context is specifically characterized in terms of the internal structure of the categories occurring in a text. In other words, the inclusion and membership properties of the semantic categories found by Rosch (1973b) is taken as an underlying principle in structuring context.

Thus, for the purposes of this study, it is assumed that reading and vocabulary depth learning are interrelated in a significant way that successful vocabulary depth learning depends on a successful understanding of a text, so that learning a lexical item means understanding its relations with other items occurring in a text. Given these considerations, this research starts with a re-definition of vocabulary depth in terms of the internal structure of the semantic categories that occur in a text. Successful vocabulary depth learning presupposes the ability to recognize these properties as they are revealed in a text.

Chapter 3   Cognitive Linguistics and Categorization

3.1   Introduction

In this chapter, we are going to explain the theoretical basis of cognitive

linguistics, especially with respect to categorization theory.   First, a distinction

between cognitive linguistics and generative linguistics is necessary, given that they

adopt very different viewpoints on the relation of language to general cognitive

abilities.   The second section will explain categorization theory and how the notion

of "prototypicality" enables us to recognize that semantic categories have internal

structure.   In particular, the research studies of Rosch will be treated as a

fundamental basis for recent theory.   The internal structure of categories, both

vertical and horizontal, will be explained.   Finally, we discuss the three levels of

category, superordinate, basic and subordinate, that are generally discussed in

experimental psychology.   As we shall see, each level has unique characteristics

that have been found in empirical studies.   We will also argue for the primacy of the

basic-level which is commonly agreed.   The reason for the comprehensive

discussion of category structure and category theory that follows is that

categorization theory underlies the current research.   Moreover, the experimental

approach adopted in this project and described in chapters 4, 5 and 6 is believed to be

largely original.   For that reason, and in the hope that subsequent researchers will

adopt and refine the present approach, the researcher considered it especially

important to provide the kind of detailed rationale that the cognitive linguistics

literature provides.

3.1.1 Cognitive Linguistics and Generative Linguistics

Cognitive linguistics adopts a very different perspective on the nature of language from generative linguistics. Three distinctive aspects are considered in this section, i.e. autonomy, modularity and innateness.

In generative linguistics, linguistic knowledge is considered as a distinct module in the human's brain. According to the modularity hypothesis, our brain is highly differentiated and each function is allocated to or managed by a distinct module or a faculty. Each faculty or module functions autonomously and is independent of the others. Thus, the language faculty is separate from other cognitive faculties, and language is processed independently of other cognitive functions. Thus, we say it is autonomous.

For cognitive linguists, however, language is placed in the context of the human being's interaction with the environment, including the objects, events and people surrounding individuals. Cognitive linguistics studies the way in which the language system is shaped by human beings' conceptualization of the world, by how they perceive it and how they interact with it. This experiential view (Ungerer & Schmid, 1996: xi), termed by Lakoff "experiential realism" (1987: xv), is firmly grounded in experience of the world. The study of language therefore cannot be separated from our unique experience and from other cognitive abilities, since it reflects our view of the world. Both the structure and the functioning of language are closely associated with general cognitive abilities. That is to say, linguistic knowledge is never autonomous because it is shaped by our world experience and thus closely related to other cognitive abilities, such as perception, emotion, categorization and reasoning (Rudzka-Ostyn, 1993; Taylor, 1995). Simply put, the distinction between cognitive and generative linguistics reflects two contrastive

perspectives: whether linguistic structure is influenced by other cognitive abilities, or whether language is a pure form, autonomous from other cognitive abilities.

Moreover, the differences between these two approaches to language study are also revealed in their perspectives on the relationship of linguistic and non-linguistic
o
or encyclopedic knowledge. Generative linguistics makes a clear distinction between linguistic and non-linguistic knowledge, given the premise that linguistic structures are autonomous and therefore necessarily separate from general conceptual knowledge. For Chomsky, linguistic *competence* and *performance* should be distinguished (Chomsky, 1965). Although the latter is a reflection of the former, it is also affected by many factors, such as affect, motivation, attitude and so on. Therefore, if we wish to study "knowledge of language" (Chomsky, 1986), it will be competence rather than performance in which we are principally interested, at least in the generative account. Generative linguists believe that treating competence as the object of scientific enquiry will result in a more coherent theory. Thus, the external language we observe serves as data which enable us to determine the nature of the cognitive structure which generates it.

In contrast, cognitive linguistics takes knowledge as stemming from our bodily interaction with the world and is thus grounded in bodily experience. Since it seeks to put language at the center of human interaction with the environment, there is no need to make a distinction between competence and performance (Rudzka-Ostyn, 1993: 1; Taylor, 1995: 18). Linguistic knowledge is considered one aspect of the conceptual knowledge shaped by our experience and thus not separable from non-linguistic knowledge. For this reason, cognitive linguistics is often criticized by generativists as the pursuit of linguistic performance (Fesmire, 1994). Seen from this perspective, the role of conceptual knowledge in language use and understanding

is a significant discussion of the cognitive approach, which is intended to provide a comprehensive understanding of how mind, body and language interact.

With respect to language acquisition, the generative approach proposes that linguistic ability is innate so that children can eventually construct a complete knowledge of a language, or a number of languages, even if they receive comparatively limited linguistic experience. The primary linguistic data to which the child has been exposed are regarded as impoverished and insufficient to account for the linguistic knowledge of the adult. The child is thus endowed with the ability to acquire his/her native language successfully. Language acquisition is accordingly regarded as biologically (genetically) rather than experientially determined (Chomsky, 1968, 1976, 1980). In contrast, cognitive linguistics claims that in order to make sense of our environment, our conceptual knowledge is derived from our real-world experience. This includes linguistic knowledge. Thus language is experientially determined (Lakoff, 1987).

Given these different fundamental beliefs about the nature of human language, cognitive linguistics adopts a different research methodology from generative linguistics. Looking for more information with respect to general cognitive abilities in order to broaden its views on human cognition, the cognitive approach looks into and incorporates studies from related disciplines, such as cognitive psychology, experimental psychology, psycholinguistics, anthropology and so on. Since the studies in these disciplines are directly related to our understanding of language and conceptual organization, they are able to enlighten cognitive linguistics both as theoretical conceptions and to the extent that they have developed empirical research strategies that may be appropriate to the study of language. This is a crucial reason that makes cognitive linguistics especially cognitive relative to generative linguistics (Gibbs, 1995: 29).

Cognitive linguists therefore assume that our cognitive abilities project on to linguistic structure. One of our most notable cognitive abilities is categorization, and, unsurprisingly, this has a significant role to play also in language. The prototype effect can serve as an example. The theory of markedness is one form of prototype effect evident in linguistic structure. The unmarked value is considered as basic and thus creates an asymmetry among category members (unmarked vs. marked), which is an effect of prototypicality as a standard by which we discriminate good and non-good members of a category (Lakoff, 1987: 59). In addition, the internal structure of categories reflects our experience in the world and, together with the primacy of basic level conceptualizations, underlies our concept of abstract notions and event categories (Rifkin, 1985). Thus, to turn things on their head, linguistic categorization can in turn inform us of our more general cognitive categorization (Lakoff, 1987: 58).

3.1.2 Categories and Prototypes

Categorization is a way of organizing the knowledge we obtain from our experience of the outside world. By grouping the objects or events we encounter into categories, we increase our ability to manage the environment around us. This "cognitive economy" saves us time and energy and enables us to store a huge amount of information regarding individual members (Rosch, 1978: 27). Bruner, Goodnow, and Austin characterize the function of categorization in the following way:

> But were we to utilize fully our capacity for registering the differences in things and to respond to each event encountered as unique, we would soon be overwhelmed by the complexity of our environment. Consider only the linguistic task of acquiring a vocabulary fully adequate to cope with the world of color differences! The resolution of this seeming paradox - the existence of discrimination capacities which, if fully used, would make us slaves to the particular - is achieved by man's capacity to categorize. To categorize is to render discriminably different things equivalent, to group the objects

and events and people around us into classes, and to respond to them in terms of their class membership rather than their uniqueness (Bruner, Goodnow, Austin, 1956: 1)

The process of categorization occurs unconsciously and poses no difficulty for normal human beings. It is the nature of human cognition, one of the most fundamental aspects of higher cognitive function.

What then is a category? Smith proposes three critical characteristics (1990: 36). First, a category allows us to code experience. Thus, we encounter an object or event, remember it as a member of a certain class and retrieve it from that class. Second, a category allows us to made an inference about new objects. Knowledge about a category represents a summary of the shared properties of its members. If we know the category name of a newly encountered object, we can predict its properties based on the category name. This "inductive inference" is an important function of categorization with which human beings are endowed at a very young age (Smith & Medin, 1981: 1). For instance, our knowledge about *mammals* tells us that whales can bear live young rather than lay eggs like fish, because a whale is a mammal. Empirically, children were found to be capable of predicting unseen properties of a category member, given the category label and some properties of other category members (Gelman & Markman, 1986). Third, similarity predicts that members of a category are similar in their overall shape or character but dissimilar from members of other categories (Smith, 1990: 36).

Instead of having necessary and sufficient attributes in a category, category members actually have few attributes in common and may resemble each other by only one or a few of the attributes found in other members. Thus, there are no common properties shared by all members of a category, rather members are linked by the notion of family resemblance, just as members of a family share one or a few features, such nose size, eyes colour or hair colour, with other members.

Wittgenstein was the first to discuss this notion of resemblance in his discussion of the meaning of *game* (Wittgenstein, 1953: 31e). While board-games, card-games, ball games and Olympic games do not have any single property in common, they still belong to the same category, i.e. *game*. It is the notion of family resemblance that ties them to the same category. Since category members do not share a single set of common properties, a natural category should be defined in terms of a set of properties, with no single one sufficient to define a category. Indeed, prototype theory reveals a universal tendency toward uncertainty with respect to category definition, and thus "vagueness" characterizes our knowledge of categories (Hampton, 1993: 70).

Rosch and Mervis (1975) tested the notion of family resemblance experimentally. Six categories, two natural ones, i.e. *fruit* and *vegetable* (natural kinds) and four artefact categories, i.e. *furniture, vehicle, weapon* and *clothing* (nominal kind) were chosen and twenty members for each category were selected. The task required subjects to list what they took to be the attributes of each member. The resulting data showed that very few attributes were shared by all twenty members of a category; in four of the categories only one common attribute was shared by all the members, whilst the other two categories, no single shared attribute was found (Rosch & Mervis, 1975: 580). The category members in this research clearly demonstrated that it is the notion of family resemblance that relates them.

In her ground-breaking research into perceptual domains, Rosch had earlier revealed new insights regarding the structure of categories (Rosch, 1973b). One of her earliest and best known experiments was conducted on Dani people whose language contained only two colour terms. Dani people were trained to learn colours which were paired with Dani clan names. (These names were chosen because they constituted a recognizable set of words unassociated with perceptual

phenomena.) Focal colours, i.e. the perceptually salient colours, were learned faster, even when hues with longer or shorter wavelengths were centrally placed by the researchers. Rosch argued that categories have internal structure with the best exemplar or exemplars being situated at the centre and other members surrounding at a distance in proportion to their similarities to these members (to use a spatial metaphor). These central members are the prototypes of the category. To put it another way, category members are defined by the degree to which they approximate to a prototype, rather than in terms of necessary and sufficient features (or criterial attributes), as claimed in the classical view. Prototypes, also referred to as the "clearest cases" (Rosch, 1973b: 112), are "the objects which most strongly reflect the attribute structure of the category as a whole" (Rosch, 1977: 3), and can be defined as "those category members to which subjects compare items when judging category membership" (Rosch & Mervis 1975: 575). However, they should not be confused with any particular exemplar, for the notion of prototypicality represents a central tendency of the category.

Furthermore, Rosch applied this finding to semantic categories including both natural and nominal kinds, each containing six exemplars (Rosch, 1973b). Natural kind categories are biological taxonomies such as *fruit, bird,* and *vegetable;* while nominal kind categories include *sport, disease, science* and so on. This experiment involved first asking subjects to rate each exemplar within its category and then ranking these exemplars (experiment 3). Rosch found a high degree of agreement as to the "best examples". A further experiment measured the reaction time to the sentence "A is a B", when A was a member and B was a category name. For example, subjects were required to verify the sentences "A carrot is a vegetable" and "A mushroom is a vegetable" (experiment 4). The result confirmed her prediction

that the reaction time was faster when A was a central or prototypical member than when it was a peripheral member (Rosch, 1973b: 138).

This research proved that members of a category do not have an equal status in cognition. Rather, some members are more typical than others, so that categories include less good members with a peripheral status within the category. The notion of degree of membership presents a great challenge to the classic view of categories, which takes exemplars as either members or non-members, i.e. in an all-or-none fashion. Besides, poor exemplars of a category may gradually merge into contrasting categories, thus entailing a fuzzy boundary of the category (Ungerer & Schmid, 1996). This also contrasts with the classic view of categories as entities with clear-cut or "absolute" boundaries (Mervis, 1980: 282).

Some researchers contrast well-defined and fuzzy categories. In one verification of category membership task, subjects were asked to make category membership judgment on pairs of exemplars and category labels in two separate sessions (McCloskey & Glucksberg, 1978). The result showed that most subjects agreed on the exemplars of high typicality as members of a category and those of low typicality as non-members of the category. However, for exemplars of medium typicality, subjects showed greater inconsistency when required to make judgements on more than one occasion. It was concluded that the perceptions of category boundaries are liable to change and thus the boundaries of natural categories tend to be fuzzy.

According to the prototypical view of category structure, membership is determined by comparing the similarity between an instance and the category's prototype (Medin & Barsalou, 1987: 463). That is to say, members of a category are actually ordered by degree of similarity along a distance dimension. Similarity is therefore argued to be the base of category membership (Hampton, 1993).

Typicality is the measure of similarity which predicts that some members will be more similar to the prototypical members of a category, i.e. more representative of the category (Baraslou, 1987; Hampton, 1993). This phenomenon in which various degrees of membership are displayed along a continuum is called graded structure (Rosch, 1973a; Barsalou, 1983: 211).

Graded structure exists not only in fuzzy categories, it is also present in well-defined categories such as *even* and *odd numbers* (Armstrong, Gleitman & Gleitman, 1983). In their research, Armstrong *et al.* asked subjects to rate exemplars of eight semantic categories, including four "well-defined" categories and four "fuzzy" categories. Surprisingly, the results showed that graded membership existed even for well-defined categories such as *odd number* and *even number*, an effect which is explained as a prototype effect. For instance, on a scale of 7-points (where *1* signified a good example and *7* a poor example), the exemplariness rating was 1.1 for *4*, and 3.9 for *106*. This means that for this group of subjects, *4* was considered a better member of the category *even number* than *106*. Thus, not all even numbers were equally good members of the category to which they belong. In the task of verifying category exemplars, reaction time also correlated with degrees of exemplariness for both fuzzy categories and well-defined categories. Since in a later task subjects stated that exemplars of these well-defined categories could not be graded, the representation of the mental concept of a category appeared to be a complex issue.

The conception of graded structure is well established in empirical work in relation to common categories, i.e. natural kinds and artefacts (Rosch, 1973b, 1975; Rosch & Mervis, 1975). It is also found in other category types (Barsalou, 1983). In the research into the structure of *ad hoc* categories, graded structure was also found in these goal-derived categories, which were formed to achieve a particular

purpose. For example, the category "ways to escape being killed by the Mafia" was provided with six 6 possible actions as exemplars, and subjects were asked to rank them (Barsalou, 1983: 215). The rankings of these test exemplars achieved a high level of agreement among subjects, thus proving that graded structure also exists in *ad hoc* or instantial categories. However, it was found that graded structure is unstable in that people appeal to different criteria in judging typicality (Barsalou, 1987: 104), and that these criteria are affected by various contexts.

To sum up, we return to Geeraerts' four characteristics of prototype theory (Geeraerts, 1989: 592-3). First, category membership is not associated with a set of necessary and sufficient criteria, as claimed in the classical view, which regards these attributes as criteria for category membership. Second, category members display a relationship of family-resemblance, as illustrated in Rosch & Mervis (1975). Third, category members do not have equal membership status in a category, but present various degrees of membership. Fourth, the boundaries of categories tend to be fuzzy, i.e. there is no clear-cut separation between members and non-members.

3.1.3 Levels of Categories

The internal structure of categories is manifest both in horizontal and vertical aspects (Rosch, 1978). Horizontally, as we have discussed, the internal structure of categories displays prototypical members in the centre of the category surrounded by other members of decreasing similarity. Vertically, categories occur at different levels based on different "levels of abstraction" or generality (Rosch, 1978: 30). Class inclusion is most readily perceived as a relation among three levels: superordinate, basic and subordinate. That is, a category is a hierarchy arranged according to the generality or specificity of objects, events or actions. The superordinate level is said to be the most general abstraction, while the subordinate

level contains items which display the most particular meanings. Simply put, categories are hierarchical in that a member of a category may also function as a superordinate category name and thus include a set of its own category members.

The basic level is assumed to be the most salient in the process of categorization because category members at this level are said to be both informative and distinct (Rosch & Mervis, 1975; Rosch, Mervis, Gray, Johnson & Boyes-Braem, 1976, Mervis, 1980; Murphy & Brownell, 1985; Rifkin, 1985; Morris & Murphy, 1990). At the superordinate level, categories are distinct but not informative because of their generalities, while subordinate categories are said to be informative but not distinct because of their specificity. The basic level is described as a "compromise" between informativeness and distinctiveness (Morris & Murphy, 1990: 407). As early as 1958, Brown raised the issue of why some words were chosen more often to name a thing than others. It seemed that some naming words had a preferred status with a higher frequency of occurrence. Despite the fact that most of these names were shorter, Brown concluded that they tended to cluster at the basic level, which he referred to as the level of "usual utility" (1958: 19). Murphy & Lassline also speculated that it is the usefulness of basic level terms that make them more favoured (Murphy & Lassline, 1997: 105).

In the literature, we can find much evidence from empirical studies that demonstrate the superiority of basic level categories. First of all, the most prominent study by Rosch *et al.* (1976) presented reliable evidence that, when asked to list attributes at a basic level of nine taxonomic categories including three natural kinds and six artefacts, subjects listed significantly more properties of basic level categories. Superordinate categories had few attributes in common; subordinate categories had more properties (but not significantly more) than basic level categories (experiment 1). Thus, the basic level was said to be more informative.

In a second experiment, subjects were asked to describe their bodily actions when they interact with category members at each level. Members of basic level categories were described with more common actions than the other two levels. In an identifying task, objects at basic levels could be identified more rapidly based on overall shapes (experiment 7). The basic level maximizes within-category resemblance while minimizing between-category resemblance. The advantage of basic level was also shown in a naming task in which subjects named the objects shown in pictures more frequently with basic level terms (experiment 10).

One may suspect that it is the frequency of occurrence or the morphological simplicity of basic level terms that renders this kind of advantage. That is, because basic level terms are typically shorter and used more often in daily life, they come to subjects' minds more readily. To exclude these possible factors, Murphy and Smith designed an experiment in which artificial categories were used (Murphy & Smith, 1982). These categories were said to be artificial because of the novel labels which were attached to familiar objects. The novel labels consisted of a CVC string (consonant-vowel-consonant) previously heard by subjects. Four basic categories were chosen, each containing two subordinate categories and were themselves grouped into two different superordinate categories. The subjects were divided into three groups and given the task of learning these three levels of categories in three different orders. The task required subjects to respond "true" or "false" to category names spelt out and followed by pictures presented by experimenters. In other words, if the category names matched the pictures, a "true" answer would be required, and if they did not, a "false" answer would be more appropriate. Consistent with the results of Rosch *et al.* (1976), basic level categories were responded to fastest regardless of learning order, thus illustrating the superiority of basic level categories over the other two levels (Murphy & Smith, 1982).

Developmentally, Rosch *et al.* (1976) also found that, when asked to sort pictures of objects at the basic level and the superordinate level, children of all ages sorted the stimuli at the basic level correctly but had more difficulty sorting those at the superordinate level (Rosch *et al.*, 1976, experiments 8 & 9). Experiment 8 required children aged 3, 4, 5, 6, 8 and 10 to pair pictures of the same kind from triads. The results showed that for the basic level, 3-year-old children achieved 99% correct answers, and 4-year-old children and above achieved correct pairings in 100% cases. However, for the superordinate level, 3-year-old children only achieved 55% rate of correct responses, while 4-year-olds achieved 96% correct answers. These findings confirmed the researchers' prediction that categorization at the basic level produced significantly more correct answers than at the superordinate level for children, and that categorization at the basic level posed no difficulty for younger children whereas superordinate level categorization was more problematic. The correct rate at the superordinate level was found to increase with age. In addition, experiment 9 set a different task which asked children to group pictures of the same kind from a set of pictures. Similar results were found. For basic level grouping, all but one child in the 5- and 6-year-old groups sorted the pictures in an adult manner; children above 6 years old performed in the same way as adults. For the superordinate level grouping, only half of the children aged 5 and 6 could sort the pictures in an adult manner, but the performance was also found to improve with age. These two experiments further established the superiority of basic level over superordinate level categories with respect to the ease of acquiring categories.

Mervis and Crisafi (1982) also showed that the acquisition of basic level categories occurred earlier than superordinate and subordinate level categories. Children of three age groups, 2.6 (2 years 6 months old), 4 and 5.6 (5 years 6 months

old), were presented with a standard object and asked to choose which one of two pictures went with the standard one. The result showed that children performed better when the standard and the correct one came from the same basic level, and the performance developed with age. 5.6-year-olds performed better (above chance level) for all three levels; 4-year-olds did well at superordinate and basic levels; and 2.6-year-olds only succeeded at the basic level. The researchers therefore concluded that the order of acquisition occurs at the basic level first, then at the superordinate level, and finally at the subordinate level (Mervis & Crisafi, 1982: 265).

Of the factors which are held to influence the primacy of the basic level, context and expertise are among those most frequently researched (Murphy & Wisniewski, 1989; Tanaka & Taylor, 1991; Lin, Murphy & Shoben, 1997). In Lin *et al.*'s experiment, subjects were presented with pictures of objects belonging to the same superordinate category and then required to think about their functions, materials, essential parts, energy sources or physical attributes of these objects (Lin *et al.*, 1997, experiment 1). It was assumed that such a task could evoke conceptualization at the superordinate level and trigger superordinate concepts. Then a categorization task followed, asking subjects to decide whether a picture was a member of a category name shown earlier. It was found that for exposed categories (those which had been shown in a prior task), the accuracy difference for superordinate and basic levels was only 5% (basic 96%, superordinate, 91%), while for unexposed categories (those which were appearing for the first time in the current task), the difference increased (basic 96%, superordinate 86%), although it was still found that responses to basic level items were faster in both cases. As a result, prior processing at the superordinate level was found to reduce the difference between the reaction times to the superordinate and to the basic levels, and thus the

superiority of the basic level was said to be reduced. It was concluded that a prior context is able to reduce the differences between the reaction times to the superordinate and basic levels (Lin *et al.*, 1997: 47).

Rosch also suggests that expertise can change the structure of categorization hierarchies (Rosch, 1977: 42-43; Rosch *et al.*, 1976: 340). Their abundant knowledge can effectively distinguish objects at a subordinate level and therefore what we consider as a subordinate level would be seen as basic level for these experts. The suggestion is confirmed in a study by Tanaka and Taylor (1991). Twelve dog and twelve bird experts were involved in the experiment. In the task of feature listing, subjects listed as many features at the subordinate level as at the basic level in their domain of expertise. Also in the task of naming, experts used subordinate level terms to name objects in their expertise domain slightly more often than basic level terms, although the difference was not significant. But for novice domains, basic level terms were used to identify objects significantly more often than subordinate level terms, confirming again the study of Rosch *et al.* (1976). The third task, that of category verification, replicated the same result. That is, subjects gave faster responses for objects at the subordinate level in their expert domain than for objects outside their expert domain, while performance at the subordinate level was the same as at the basic level in their expert domain. Based on these results, it was suggested that "basic level" could be redefined for different individuals as a consequence of expertise (Tanaka & Taylor, 1991: 477). What is considered to be basic level therefore appears to vary with personal domain knowledge. This supports Mervis, who suggested that basic level should be seen as relative rather than fixed (Mervis, 1980: 291).

Finally, some researchers have explored the extent to which fundamental differences may be said to exist between basic and superordinate level terms.

Wisniewski & Murphy (1989) analyzed superordinate and basic categories present in texts chosen from a large body of prose of various styles and subjects, and found that superordinate categories occur more often with multiple objects, while basic categories always referred to single objects (Wisniewski & Murphy, 1989: 255). This was said to be due to the fact that basic level terms are informative and distinctive, and used to designate individual objects and provide as many characteristic properties as possible as a means of singling out individuality. In contrast, superordinate terms are general and leave out many unique characteristics of objects; thus they are used as a cover term for a group of objects with common properties.

In a follow-up study, Murphy & Wisniewski found the difference between verification of category membership at superordinate and basic levels was reduced if objects were placed in their natural environment, i.e. an environment constituted by other members at the same level (e.g. three kinds of vehicle) located in an expectable configuration (e.g. traveling along a road). The subjects were asked to judge whether an object placed in isolation or in a "scene" as a member of a category name previously indicated. It showed that an object presented in isolation was more readily recognized a) at the basic level and b) as a member of a superordinate category than when it was presented in a scene. However, the differences between the times taken to recognize an object and to assign it to a superordinate category were less when the object was resented in a scene. They argued that a superordinate category contained relational information about category members, and its scene helped the subjects to recognize their relations between members at the superordinate level, but not at the basic lever (Murphy & Wisniewski, 1989: 576). A functional difference is therefore suggested as another fundamental difference between superordinate and basic level terms (Lassaline, Wisniewski & Medin, 1992: 366). Additionally,

Cruse suggested that superordinate and basic level terms have different pragmatic functions. When intending to be vague, one can use superordinate terms to label a particular object, thus obscuring its characteristics (1977: 163). Such research clearly illustrates that superordinate and basic terms are very different conceptually.

3.2. Meanings of Words

3.2.1 Words and Concepts

Words are not concepts but serve to activate concepts in our minds. As Clark says, "words flag concepts" (Clark, 1983: 797). That is, a word evokes a mental representation in relation to its meaning. A word meaning is not exactly the same as a concept, but built out of a concept (Murphy, 1991). A concept is the mental base for word meanings and comes into existence earlier than word meanings. Words are linguistic expressions which are conventionalized by social communication and cultural background shared by the people within the same community. Almost always, we characterize the meanings of words in terms of the relation between concepts, which in turn are based on our experiences (Lakoff, 1987: 207). For example, Eskimos have different words designated for different types of snow, which do not exist in other languages. The existence of words is thus determined by our experiential conceptualization.

On the other hand, concepts cannot be independent of language. Through language, concepts become tangible (Vygotskii, 1986: 218). Besides, our concepts are formed by the information and knowledge that we receive every day. To obtain more information or to pass on knowledge, we mostly rely on language. As such, concepts cannot have continuous existence without the aid of language. But not all concepts are lexicalized or are capable of being expressed by means of words. Words are only a subset of concepts although there is some "overlapping" (Aitchison,

1987: 41) between words and concepts.   Concepts largely exist beyond the extent to which words can denote them.

To investigate the meaning of a word, early linguists followed philosophers in distinguishing the notions "sense" and "reference".   Sense refers to the concept a word is associated with while reference refers to the set of things a word designates. Sometimes, the reference of different words or phrases may be the same, while their senses are different.   A much quoted example is Frege's observation that Venus can be referred to as either the "evening star" or the " morning star" (Geach & Black, 1980: 57).   The reference of a word can vary to the extent to which individuals conceptualize it via their unique experience.   What we call *desk* may have different sizes and shapes in Western countries and ancient China, though desks have the same function wherever they are found.   On the other hand, different linguistic expressions, inter-lingually or intra-lingually, can have the same sense.   It is also possible that words may have no reference in the real world, but still have some sense.   All these possibilities indicate that sense exists in our conceptual system and approximates the notion of concept.   Therefore, to define the meaning of a word, we usually appeal to the concept it represents (Cruse, 1986: 49).   According to Saeed, sense "places a new level between words and the world: a level of mental representation." (1997: 32). Words are meaningful only in our conceptual system. Words denote concepts and concepts substantiate word meanings.

This discussion reveals the mapping problem between words and concepts. The mapping patterns between words and concepts can vary.   When we encounter a word, it immediately calls up a related concept.   The simplest case is that a word entirely maps onto a single concept.   This kind of one-to-one relationship happens rarely.   Sometimes many words can denote the same concept.   This occurs when we encounter synonyms, though there are very few exact synonyms.   It is also

possible for a word to be connected with multiple concepts, i.e. a homonym is able to cause this effect if it is surrounded by an ambiguous context and cannot be constrained by context (Murphy, 1991).

Nevertheless, the structure of our internal lexicon is always thought to mirror that of concepts. Historically, for most cases, words and their conceptual counterparts are assumed to exist in a connected network in our memory (Collins & Quillian, 1969). Concepts are stored as interconnected links and the network displays a hierarchical structure consisting of nodes linked to each other and representing sets of properties (1969: 241). Once a word is encountered, it activates related concepts and their properties. Since a word may call up multiple concepts, context is necessary to constrain the possible interpretations of a word. The effect of context plays a crucial role in the process of pinning down the meanings of words.

3.2.2 Concepts and Category

Categories organize our concepts. As indicated in the first section of this chapter, the process of categorization, i.e. that of assigning objects and events to categories, is fundamental to cognition. It is the way we handle the massive quantity of information we receive from the outside world. The process of categorization helps us determine whether or not an object or event can be classified as an instance of a particular category; and more than that, it also helps us predict unperceived properties possibly present in that category. In other words, the conceptualization of categories gives us a clue for inferring additional properties in a particular instance (Smith & Medin, 1981: 9; Smith, 1995: 5-6). This kind of mental process is no less important than categorization, for it equips us with the ability to take in a large amount of information efficiently on the one hand, and thus reduce our effort at further categorization on the other.

Moreover, category membership can also be used to determine the relationship between concepts. This benefit is most conspicuous when we talk about the hierarchical levels of categories, especially the taxonomic hierarchies denoted by concrete noun objects. The conceptual hierarchy representing category relation is arranged in such a way that, when we encounter an instance, it enables us to access a concept and retrieve related properties of the instance. In addition, the primacy of the basic level in the taxonomy hierarchy, as discussed in the section 3.1.3, provides a credible explanation of how we acquire concepts.

In a word, concepts are the mental representations of categories stored in our memory (Smith, 1995). To define a concept, the classical view of conceptual representation states that a concept contains properties that are "singly necessary" and "jointly sufficient" (Smith & Medin, 1981: 23). These properties are viewed as "defining features", which we use to determine whether an entity is an instance of a category concept. In other words, to be an instance of a concept, these properties must be present. Nevertheless, for the prototypical view, there are no "defining features" but only characteristic properties.

Given the proposition that words can activate concepts, and that concepts are structured in categories, we can find a similar relationship between word and category. Just as there is no straightforward relation between word and concept, there is no one-to-one relation between word and category. More problematic is cross-classification, i.e. when a word can call up more than one category in our minds, each one representing a different possible categorization based on different criteria or functions. A *dog*, for example, can be classified as an instance of the category *animal* or *pet*. The word *green* can remind us of the colour category or of a political connotation. These candidate interpretations require a specific context for an appropriate meaning to be determined.

3.2.3 Context

The interaction between environment and human beings are what cognitive linguists are interested in for they have the potential to shape our conceptual knowledge and hence its linguistic representation. Therefore, in this section, we will consider the way in which context is able to affect the meanings and interpretations of linguistic elements. Many studies in psycholinguistics as well as in cognitive and experimental psychology that directly bear on the effect of context in the determination of lexical meaning and category structure will be drawn on. As one of the methodological concerns of cognitive linguistics is to broaden its research methods, this survey of empirical studies from related disciplines has a two-fold benefit. The findings obtained in these studies are treated as the basis of theorizing and the methods employed are worthy of consideration as possible research strategies in cognitive linguistics.

3.2.3.1 Word Context

When we discussed the notion of context in our discussion of vocabulary learning in chapter 2, we were referring to the linguistic context in which a particular lexical item occurred. As has been indicated in that chapter, context provides clues which enable the learner to infer the meanings of unknown items. In addition, context can help a reader determine the meanings of words, especially in the case of polysemies and homonyms. Because of its role in enabling us to infer the meanings of words, context is essential for successful comprehension in that it establishes a relevant semantic environment which enables the retrieval of lexical meanings.

In effect, reading comprehension involves various levels of processing. The first step is to recognize a string of letters as a representative of the phonological

matrix of a word and then to retrieve the meaning of that word from memory. The retrieval of meaning may be problematic if the word is unknown or may be delayed if an ambiguity occurs. The context constituted by surrounding words may provide clues for guessing the meaning of a particular item. Context, in other words, can help to relate all the given words to form a coherent and meaningful sentence and further to relate all the sentences so that the message context of a text may be recovered. We can therefore say that the appropriate understanding of a text depends to a considerable degree on context.

Therefore, in this section the effects of context will be discussed in relation to different levels of processing, i.e. word, sentence and text. These are all influential in the retrieval and determination of lexical meanings or in determining the interpretations of referents intended in a sentence or a text. The discussion will be sequenced in the order from basic to higher levels of processing involved in understanding a text. That is, the studies of recognition of words will be reviewed first so as to show how context can speed up or slow down the process of recognizing a string of letters as the phonological matrix of a word. This will be followed by a discussion of how context can resolve an ambiguity occurring in a sentence and determine an appropriate meaning. Finally, we will discuss how the wider context of a whole text makes salient a particular property or set of properties of a lexical item and how this wider context helps a reader to select a particular item as the referent of a lexical item so as to enhance reading comprehension.

The semantic environment constituted by single words will be discussed first. Earlier research usually grouped word stimuli into pairs and presented them to subjects whose task was to make a decision as to whether a target word (i.e. one of the two words) was an English word or a non-word (Meyer, Schvaneveldt, & Ruddy, 1975). In a pair, the first word might be semantically related or unrelated to the

second word, for example BREAD-BUTTER and NURSE-DOCTOR, or NURSE-BUTTER and BREAD-DOCTOR (1975: 99). The reaction time to each word was measured and compared. It was found that the response to the second word was facilitated if it was semantically related to the first one. The effect of context was argued to rely on the environment established by the first word in this type of word pair and therefore enhance the recognition of the second word (Meyer *et al.*, 1975: 100).

To expand this paradigm, Schvaneveldt, Meyer, and Becker designed a different experiment in which three, rather than two words (i.e. word triplets), were presented one by one with the second item a homograph, i.e. a word having two or more distinct meanings (1976). Different types of linguistic contexts were specially designed, depending on the semantic relations between the first and the second words, and between the second and the third words. The first word might be either related or unrelated to the second one, and similarly the second might be either related or unrelated to the third. In a more interesting design, the first word might be related to one meaning of the second and the third word might be related to the same meaning or to the other meaning of the second since the second word was a homograph. For example, the first and the third word are related to the same meaning of the second word as in the word triplet "SAVE-BANK-MONEY", and are related to different meanings of the second word as in "RIVER-BANK-MONEY" (Schvaneveldt *et al.*, 1976: 246). Subjects were required to decide whether or not each word was an English word, and the reaction time for each was measured.

Apart from the expected result that the response to the related words relative to unrelated ones was faster, it was demonstrated that the reaction to the third word was much faster when the first and third words were related to the same meaning of the second word than when the first and third words were related to different meanings

of the second word. The reaction to this type of context containing divergent lexical association (e.g. "RIVER-BANK-MONEY") was shown to be slowest, even slower than when the third word was unrelated to the second one, as in the example "FIG-DATE-MONEY" (Schvaneveldt *et al.*, 1976: 249). The researchers explained these results by suggesting that, because the first word had activated certain features related to one meaning of the second word, it provided a positive semantic environment which helped access the concept of the second word, but it could become highly constraining in making judgements about the third word if the third word was related to a different meaning of the second word. The effect of context was thus proved to be especially strong because it was shown to cause interference in the recognition of the third word in the sequence.

This research showed that the recognition of a string of letters as the phonological matrix of a word can be either speeded up or slowed down by the semantic relatedness or incongruity of surrounding words. In a related context, the recognition process is enhanced but in an incongruous context the process is impeded. This suggests that in reading comprehension, the level of word recognition is likely to be affected by its semantic environment.

Furthermore, sentence context can also facilitate word recognition. This was demonstrated in the research of Schuberth & Eimas (1977). Three types of context were provided for target words: isolation, sentential context and digital context. Isolation meant no context was provided; sentential context involved putting target words in a sentence; and digital context referred to four randomly selected digits placed together before the appearance of target words. After seeing a type of context, subjects were asked to make a decision on the acceptability of a target word. Generally speaking, the results accorded with the prediction that sentential context could facilitate word recognition relative to no context and digital context.

Moreover, the degree of congruity was differentiated among target words. The congruity of target words was defined in terms of their appropriateness in their sentences and further ranked by a different group of subjects. For example, in the sentence "THE PUPPY CHEWED THE _____", BONE would be more congruous than HOUR in the blank space (Schuberth & Eimas, 1977: 29). As the results showed, the reaction times to congruous words were faster than to incongruous words of the same frequency level because sentential context made congruous words more accessible and restricted the processing of incongruous words. The effect of context was more apparent when comparing the reaction times to incongruous words in sentential contexts than when that word appeared in isolation, in that the interference of sentential context slowed down the recognition process. This implies that not only can context enhance word recognition, but that it can also constrain the range of possible meanings. In other words, the context of a text sets limits on the potential words that can appear. This kind of semantic constraint constituted by a context is necessary to make more effective the retrieval of word meanings from semantic memory.

It was also found that single words might contribute to disambiguating homographs in a sentence. Vu, Kellas & Paul (1998) specifically chose a set of verbs (experiment 1) and nouns (experiment 2) to test whether they could function to disambiguate word meanings within a sentential context. These verbs and nouns were used in simple sentences to prime related meanings and were compared with the effects of general verbs and nouns. Three sentence types were ambiguous, dominant biased or subordinate biased as in 1) "He located the bat"; 2) "He splintered the bat"; or 3) "He wounded the bat" (Vu *et al.*, 1998: 981). The verb "splintered" was related to the dominant meaning (more frequent) of "bat", while "wounded" was related to the subordinate meaning (less frequent). Both were

supposed to be more specific than the general verb "located". Target words were also distinguished as related to the dominant (such as "wooden") or the subordinate meaning (such as "fly") of the homographs, and to unrelated words (i.e. "safe" & "station"). The naming task required the subject to name the target words as soon as possible after seeing the priming sentences. The naming latencies were measured and compared.

As was expected, reaction to dominant related targets primed by a dominant sentence type was faster than to dominant related targets primed by an ambiguous sentence type. Similarly, reaction to subordinate related targets primed by a subordinate sentence type was faster than to subordinate related targets primed by an ambiguous sentence type. A more noticeable effect of context was shown in the interference of related targets primed by inappropriate context. That is, the response to dominant related targets when primed by the subordinate sentence type was slower than when primed by the ambiguous sentence type, and similarly the response to subordinate related targets when primed by the dominant sentence type was slower than when primed by the ambiguous type (Vu *et al.*, 1998: 983). Thus, these results confirmed the researchers' hypothesis that a single lexical item could function to disambiguate word meanings. From this research, we can see that since related lexical items can function to disambiguate the intended meaning of a potentially ambiguous word or even to enable the subjects to determine the meaning of a hitherto unknown word, lexical clues of this kind provided in the wider sentence are a significant means that can be appealed to in understanding a message.

Not only can context make the retrieval of word meanings easier, but context can also prime relevant properties so as to render a coherent and integrated interpretation. In an earlier study by Barclay, Bransford, Franks, McCarrell and Nitsch (1974), it was found that contextually appropriate cues could work more

effectively for recall. Compare the sentences: 1) "The man lifted the piano," and 2) "The man tuned the piano" (Barclay *et al.*, 1974: 473). When the cueing phrase was "something heavy", it was considered to be more appropriate for sentence 1), than for sentence 2), where "something with a nice sound" would be more appropriate. In recalling the same target noun (i.e. piano) embedded in a sentence, the appropriateness of the cueing phrase (to a particular sentence) made a great difference in performance. It was concluded that, when the sentential context constrained the interpretation of a target word to a certain property, the cue appropriate to that property could produce better recall. In other words, this kind of context encodes a particular aspect or property of a word and helps to maintain the memory trace in long-term memory for a longer period.

More evidence was found in Anderson and Ortony (1975). They showed that context could emphasize an aspect of an unambiguous word and thus affect its interpretation and the construction of a mental representation for a sentence. Sentential contexts such as "Nurses are often beautiful," and "Nurses have to be licensed," were employed (1975: 171). When the cue word was "actress", the first sentence would be the target sentence because the word "actress" was more relevant to the aspect emphasized by this sentence. If the cue word was "doctor", the second sentence would be the target because the word "doctor" was more relevant. Although the subject noun was the same in these two sentences, different contexts in which the subject nouns occurred resulted in different interpretations. The experimental result showed that, when priming cues aroused the relevant properties of subject nouns emphasized by a particular sentence context, the recall of the target sentences was found to be better. Based on this finding, the researchers explained that during the process of comprehension, context could construct a "particularized and elaborated mental representation" by selectively emphasizing relevant properties

and building up a more detailed image (1975: 167). This research showed how context helps us to understand a message in a more concrete way so that the comprehension of this message will be more successful.

When the context was expanded to become a paragraph, the results were almost the same (McKoon & Ratcliff, 1988). After reading a paragraph, subjects received a concluding sentence and were required to make a true or false judgement. Three types of concluding sentences were provided: a matching sentence which corresponded to the properties mentioned in the paragraph, a mismatching sentence which revealed other properties, and a filler sentence which was only relevant to general knowledge. The measurement of response times showed that matching test sentences were responded to significantly faster than mismatching sentences. Moreover, the rate of accuracy in verifying matching sentences was also higher than that of verifying mismatches. The researchers concluded that context could affect the encoding of linguistic elements during the process of comprehension so that matching sentences containing the target properties were responded to faster. This research as well as Barclay *et al.* (1974) and Anderson & Ortony (1975) reveals that the encoding of a linguistic message heavily relies on its context and therefore, the decoding or recovery of this message needs to appeal to its context.

One way of understanding a message in a more concrete way is by means of instantiation. Context can instantiate a more general term so as to provide an actual image. Anderson, Pichert, Goetz, Schallert, Stevens and Trollip (1976) compared the cueing effects of low associate and high associate words with respect to a general noun (experiment 3). As defined, high associate words were better members of a category in isolation than low associate words and thus were assumed to be recalled with more ease. However, in an appropriate context, low associates proved to be better cues than high associates. For example, in the context "The nurse pointed to

the bird on the platter", "Turkey" was a better cue than "Robin" though it was not as good a member of the category *bird* as robin (Anderson *et al.*, 1976: 677). This study suggests that, when encoding a general noun, which is a category label, context helps to instantiate the noun with a specific exemplar in accordance with the context. Thus, a lexical item which is a category label can be understood as a member of the category, and renders a particular understanding of the general term and hence of the whole message.

The same results were also obtained by McKoon and Ratcliff (1989). Using the same paradigm, they designed two types of sentence contexts, predicting and neutral, both containing a category label. Two different target sentences were presented, containing either primary or secondary exemplars of a category. For example, for the category "fruit", the primary exemplar would be "orange" in the predicting sentence "The young attorney wanted to make sure she had fresh juice for breakfast, so she bought and squeezed the fruit herself", but "grapefruit" would be considered secondary for the same sentence (McKoon & Ratcliff, 1989: 1135). The reading times for target sentences were measured and found to be faster for those containing primary exemplars than for those containing secondary exemplars. It was found that the predicting context could facilitate the reading times for target sentences with primary exemplars, but not for secondary exemplars. Therefore, it was argued that the understanding of a category term was constrained by its context so that the best exemplar that conformed to the context would be understood more readily. In other words, the exemplar that could be encoded was determined by its context which defined the properties of the category. From the perspective of reading comprehension, understanding of a category term varies from context to context because the content or properties of the category are defined by its context.

In this section, we have found consistent evidence that context can facilitate the recognition of lexical items and the retrieval of relevant meanings. Moreover, context can help to establish a more concrete mental representation by activating relevant properties or instantiating a particular exemplar, thus making comprehension easier. Based on these studies, the importance of context for reading comprehension cannot be overlooked. In Langacker's words, no linguistic phenomena should be taken as "fixed structures" (Langacker, 1997: 237), but be viewed as flexible because their context necessitates modification. In this sense, context modulates the meaning of a lexical item for an appropriate understanding. Although the notion of context employed in the above discussion is limited to that provided by single sentence exemplars and hence become much narrower than will be investigated in this research into the broader discourse context, both function as building blocks to a higher level of context, i.e. the mental representation of cognitive models, which will be considered in the current study. Only after the literal comprehension of a linguistic message has been successfully achieved in working memory, will it be possible to invoke relevant cognitive models shared in long-term memory.

To conclude this section, the effect of context contradicts the autonomous view of lexical access, thus supporting the interactive view of lexical processing that contextual information is crucial in determining the meaning of a lexical item. This constitutes evidence for the view of cognitive linguistics that there is much interaction between language and our conceptualization of the world we experience.

3.2.3.2 Context and Category Structure

Although context is usually referred to as the environment in which a word or a sentence is embedded, in the previous subsection we tended to broaden the definition

to a kind of "cognitive representation" of mental concepts (Ungerer & Schmid, 1996: 47). In this subsection, we are concerned with how context affects the category structure pre-stored in our knowledge system. It is assumed that context exerts its effects on category members by reshuffling category membership and on category boundaries by adjusting the criteria used to judge category membership. The prototypical members or best exemplars may as a result be downgraded as members of a category and the category boundaries can be extended or restricted.

As early as 1973, Labov discovered the effect of context on categorization behaviour. Twenty cups with various ratios of width to depth and different shapes were presented to subjects. Four different contexts were created in which these cups were introduced: neutral (i.e. no context), coffee, food and flower contexts. Perceiving the different contexts, the subjects named these cups as "cups", "bowls" or "vases". For instance, in a food context with cups containing mashed potatoes, more subjects named them as bowls when the width ratio reached a certain degree than in the neutral context, and likewise in a flower context in which the objects contained flowers they were named as vases (Labov, 1973: 357). However, in a coffee context where a man was drinking coffee from the cups, more subjects classified them as cups. What this experiment showed is that context can influence our perception of entities and thus the membership status of these entities may change. In other words, the inclusion of entities as category members can be a function of context. When the context changes, an object previously seen as a non-member can be included as a category member because it conforms to the properties or content of the category in a new context.

In another earlier experiment with respect to similarity judgement, Tversky asked subjects to group four countries Australia, Sweden, Poland and Hungary into two sets (Tversky, 1977). Subjects were asked to compare perceived similarities

among the stimuli. When one of the stimuli was changed or replaced (e.g. Poland was replaced by Norway), the pairing patterns were changed. It was argued that the common attributes under consideration changed and thus caused different pairings. In Tversky's words, "context" was a "frame of reference" (1977: 340). The salience of features varied in a frame, depending on what objects were available. As the to-be-compared set of objects altered, the salience of certain features, the "diagnostic values" (1977: 340), or the "classificatory significance" also changed (Tversky, 1977: 342; Tversky & Gati, 1978: 90). Tversky's work demonstrated that context can alter the classification criteria which were abstracted from an available candidate set. In this simple experiment, the inclusion properties of a category were apparently modified by a context consisting of different available examples. Again, although the notion of context in these experiments is in a different level from what would be intended to invoke in this research, it is the strict constraint on the availability of instances imposed by such a context that resembles the investigation of the contextually determined semantic categories investigated in this research.

Recognition of category membership is also affected by the physical properties of a context. As mentioned above (see p. 80), Murphy & Wisniewski's research demonstrated that an object placed in a scene required more time for identification as a category member than when presented in isolation and, consequently the effect could reduce the superiority of basic over superordinate categories with respect to their verification as category members (Murphy & Wisniewski, 1989). It was argued that a scene containing other objects at the superordinate level could interfere with the recognition of a target member because it demanded that attention be paid to other objects. While the scene of a superordinate concept made conspicuous the relational information about category members and thus facilitated recognition of

category members at the superordinate level, the scene of a basic category did not function in the same way. The results of this study suggest that the physical environment of objects can affect recognition of category members.

The mental representation shaped by prior processing can also influence the recognition of category membership. As has been mentioned (see p. 78), Lin *et al.*'s experiment led subjects to engage in thinking about the properties of the objects belonging to the same superordinate category (Lin *et al.*, 1997). It was supposed to evoke conceptualisation at the superordinate level. A subsequent task of verifying category membership at the superordinate level was then facilitated and the differences between the reaction times to the superordinate and to the basic levels were reduced (Lin *et al.*, 1997). This study accords with Murphy & Wisniewski's study (1989), in that context can diminish the superiority of basic over superordinate categories in verification of category membership.

In addition, context can also change the internal structure of a category by altering the distribution of category members within a category. This finding is derived from the series of experiments conducted by Roth and Shoben (1983), in which the relation between an exemplar and a category was manipulated via linguistic context. In the first experiment, three different types of sentence context containing a category term were presented, one of which could be biased toward a particular exemplar (biased+, e.g. "Stacy volunteered to milk the animal whenever she visited the farm"), another toward no particular exemplar (neutral, e.g. "Dorothy thought about the animal"), and the third biased toward another exemplar (biased-, e.g. "Fran pleaded with her father to let her ride the animal") (Roth & Shoben, 1983: 354). After this, subjects read a target sentence containing a particular exemplar (e.g., "She was very fond of the cow") and the reading time was measured. As had been predicted, the reading times in a biased+ context were shorter than in a neutral

or biased– context, and the reading times in biased– contexts were longest. That is, a biased+ context enabled the target sentence to be more readily understood because the meaning of the exemplar was identical or close to the meaning of the category term in the pervious sentence. It was concluded that context could change the mental representation for a category term and thus affect the ease with which potential exemplars could be retrieved.

A second experiment investigated whether or not the degree of representativeness of an exemplar within a category could be changed by context (experiment 2a). Two different sentences containing the same category term and suggesting the same best exemplar were introduced. Two target exemplars were selected. One of the exemplars (similar target) was more congruent with one of the contexts (Congruent context), but incongruent with the other (Incongruent context). The other exemplar (dissimilar target) was presented in a reversed condition. A group of subjects was asked to rate the degree of representativeness of an exemplar within a particular context (goodness-of-example ratings). The ratings were found to vary with context. For similar exemplars, they were considered more typical in a congruent context and therefore rated as more representative. In contrast, they were considered less representative in an incongruent context than dissimilar exemplars. The reverse was true for dissimilar exemplars. This demonstrated that context was able to alter the degree of representativeness of an exemplar within a category. In their words, the context can "restructure" a category by changing the degrees of representativeness of category members within a category (Roth & Shoben, 1983: 363).

Subsequently, target sentences containing similar and dissimilar exemplars were presented and reading times were measured (experiment 2b). As expected, the reading times were faster for similar target sentences following a congruent context

and for dissimilar target sentences following an incongruent context, but slower for similar exemplars following an incongruent context and for dissimilar exemplars following a congruent context. These results further confirmed the previous finding that the GOE (goodness-of-example) distribution of a category was restructured by the context (Roth & Shoben, 1983).

This series of experiments also show that context can change the internal structure of a category by altering the degree to which an item is perceived as a good member of the category. As a consequence, the ease of the retrieval of category members when presented in context differs from when they are in isolation. As demonstrated in experiment 1 of Roth & Shoben (1983), the identification of category membership was easier in context when the context strongly suggested a particular exemplar, and was difficult when the context implicitly indicated a different exemplar. In experiment 2, it was shown unequivocally in the GOE ratings that the degree of representativeness of a category member was in effect changed by a context and thus the typicality ordering of category members within a category was rearranged. In sum, the degree of category membership for a particular member is not invariant as shown when it is in isolation but varies from situation to situation, or from context to context.

A stronger degree of context dependence for categorization was found to be psychologically real in one of the research studies mentioned in section 3.2.1.2. *Ad hoc* categories were formed with the premise that, in a given context, categories can always be formed to achieve certain goals (Barsalou, 1983). For example, for a newcomer to make friends in a strange city, six ways to make friends were classified and ranked as possible members/non-members. Although these categories were not well established as common categories in our conceptualisation, they behaved like other categories by having graded structures and unclear cases (1983: 216). In

other words, *ad hoc* categories are genuine categories and their dependence on context provides further evidence that category structure can vary with context.

Given that category structure can be greatly affected by context, concepts or conceptual structure are also subject to change as a result of context. Research into size comparison between animal terms found that the size concept for animals is not fixed, but greatly influenced by the terms available and the on-going task. In their research, Cech and Shoben (1985) used twelve animal terms, ranging from *mice* to *elephant*. Generally speaking, when only terms denoting relatively small animals were involved in the comparison, the reaction times to the task "choose smaller" were faster than to the task "choose larger". On the contrary, when only terms denoting relatively large animals were available, the reaction times to the task "choose larger" were faster. Nevertheless, for pairs of intermediate size animals such as "rabbit-beaver", a different outcome appeared. When the whole set of animal terms was involved, the task of judging smaller animals was faster because this pair became relatively smaller. But when only smaller animal terms were involved, this pair became relatively larger and then the response time to the task "choose larger" was faster. This implies that concepts can be relative, and not absolute.

It has been shown that context can affect the inclusion of a category (Labov, 1973; Tversky, 1977; Tversky & Gati, 1978), the identification of category members (Murphy & Wisniewski, 1989; Lin *et al*, 1997), and the degree of category membership (Roth & Shoben, 1983). Since objects do not appear in isolation in the real world, the context or the environment in which they are encountered is always liable to affect our perception of target entities (Biederman, 1972: 78). A congruent or more natural context may enhance our recognition of certain objects, while, in contrast, an incongruent or unnatural scene may cause hindrance (Palmer, 1975).

Rosch also reminds us of the possible effects of context on the levels of categorization and on specific instances (Rosch, 1978: 42). More recently, Barsalou argues that we cannot get away from context for we always categorize to achieve some purpose in daily life, and our goals determine how categories are formed. It is only researchers into categorization as an abstract notion who would place categories in isolation (Barsalou, 1991: 45). The advantage of the chameleon-like nature of the category concept lies in its flexibility for new concept formation.

Finally, I would like to conclude by stressing that, the structure of a concept should not be taken as "invariant" (Barsalou, 1987: 114), and the mental representation of a concept should not be seen as "static" (Langacker, 1997: 237). Rather, both are subject to change as a result of context. Barsalou provides a more constructive view that concepts can be viewed as constructs in working memory which use information stored in long-term memory (Barsalou, 1987: 120). In different contexts, various properties of the current context will be transferred to working memory to build up a useful mental representation for an immediate purpose. In other words, a "concept" does not necessarily imply a fixed, well-defined representation. On the contrary, it may be context-dependent. As spelt out by Langacker, "concept", "representation" and "mental structure" need to be acknowledged as notions which are flexible. This perspective acknowledges the changing nature of our environment in which abundant interaction and communication between individuals and their environment occurs (Langacker, 1997: 232).

3.3. Overview of the Current Research

In this section, attention will be drawn to two important aspects of the current study that differentiate it from the existing L2 vocabulary learning research and relate it to cognitive studies regarding categorization. First, the notion of "vocabulary" will be re-defined with respect to the internal structure of contextually determined semantic categories that occur in a text. This definition is original and needs to be emphasized before further discussion so as to make a clear distinction between the notions of "vocabulary learning" employed by previous L2 studies and the current research. Secondly, the instantial categories investigated in this study are contextually determined and thus differ from object categories studied in experimental and cognitive psychology. In the literature, very few studies focus on this type of category. However, the investigation of categories of this kind has the potential to broaden the scope of categorization theory, and to relate object categories as studied in the most of Rosch and others, and instantial categories of the kind captured by the notion of Idealized Cognitive Models (ICMs).

3.3.1 Re-definition of Vocabulary Depth Knowledge

In this study, we define the notion of "vocabulary depth knowledge" differently from the way it is defined in the vocabulary learning studies reviewed in chapter 2. While previous studies define "vocabulary" as single word units, and "vocabulary learning" as the learning of the meaning of an unfamiliar or unknown word, or of a new meaning of a known word in a new context, "vocabulary" as investigated in this research will be concerned with multi-words items or phrases. The literal meanings of these items are supposed to be familiar, or known to informants of higher proficiency, but the learning of the cognitive categories designated by these items requires the ability to recognize the internal structure of

these categories, namely their inclusion and membership properties. Therefore, the ability to recognize these properties will be regarded as an indication of successful learning of vocabulary depth knowledge. This constitutes a new definition of vocabulary depth knowledge which has not been acknowledged in the existing literature to date.

In other words, context is characterized in terms of the inclusion and membership properties of the contextually determined semantic categories occurring in a text, and thus the meaning of a lexical item depends on its relations with others within the same category. Naturally, these relations can vary with different contexts. Successful vocabulary learning is therefore defined in this research as the ability to recognize these properties. That is, the superordinate categories need to be recognized, and lexical items need to be understood as included as more or less prototypical members under a superordinate name present in or inferred from the text. As indicated in chapter 2, no principled definition of the precise nature of context was ever offered in the L2 vocabulary learning research. Our re-definition of vocabulary dept knowledge can structure and organize context in a principled way and grant flexibility to the changing nature of contexts.

Specifically, it is the hierarchical relations that connect items in the text that accounts for successful reading comprehension. As typically understood in the literature, reading for comprehension and reading for vocabulary learning will likely compete with each other. Nevertheless, this is not necessarily because it will be shown that vocabulary knowledge can only be obtained by means of effective reading comprehension. To echo Paribakht & Wesche, reading comprehension and vocabulary knowledge are "symbiotic" (1999: 196).

While surveying the literature of L2 vocabulary learning research, a number of problems exist in the L2 vocabulary learning research with respect to the definition

of "vocabulary learning", as discussed in Bogaards (2001). Most research assumes that vocabulary learning concerns the acquisition of new words (Li, 1988; Hulstijn, 1992; Luppescu & Day, 1993; Knight, 1994; Ellis, 1995; Hulstijn *et al.*, 1996). This perspective views vocabulary as "words", and hence that the acquisition of new words means the acquisition of new forms. This point of view greatly limits the scope of investigation of vocabulary knowledge because several problems arise from such a conception. In the first place, we have to question the assumption that vocabulary acquisition means the learning of lexical units as single words. For second/foreign language learning, it has been noted that learning lexical units consisting of multiple words or word combinations such as idioms and institutionalized expressions is equally important (Nattinger & DeCarrico, 1992; Lewis, 1993; Bogaards, 2001). In this research, it will be shown that category names and category members are in fact most usually phrases rather than single words.

Secondly, learning a word is always limited to learning one meaning of a word. As is widely agreed, most lexical items have more than one meaning. That is, polysemies abound in the lexicon. A serious drawback inherent in the research on L2 vocabulary learning is that it only demonstrates the retention of a single meaning of a word in a particular text and takes successful retention as acquisition of a word (Bogaards, 2001: 322). However, as the research reported in this chapter makes clear, learners who remember only one meaning of a word are unlikely to be able to use the word in other contexts. As we know, it takes a long time and requires several phases to learn different meanings of a word and to recognize the kinds of meanings that different contexts bring about. Very often learners learn different meanings in different contexts and thus acquire vocabulary knowledge gradually and in a step-by-step manner. Therefore, it is very unlikely that language learners, and

especially foreign language learners, can learn all the meanings of a word from a single context.

To summarize all the arguments, the big question is whether vocabulary learning concerns principally the learning of new forms, or the learning of new meanings for an already known form, or the learning of different aspects of word knowledge. The existing research has tended to focus on the first two at the expense of the third of these notions. In contrast, drawing on the insights of cognitive linguistics, this research investigates the ability of learners to recognize contextually defined meanings by re-defining the notion of vocabulary depth knowledge in terms of the ability to recognize contextually determined semantic categories.

3.3.2 Instantial Categories (Non-conventionalized Categories)

Contextually determined semantic categories are different from those usually investigated in experimental and cognitive psychology. The categories investigated in this study are text-specific and non-conventionalized. In other words, they are instantial because they exist in the instances occurring in a particular text. Previous research into categorization was principally concerned with object categories, or taxonomic categories, such as *fruit* and *vegetable*. These are conventionalized and well-established in our long-term memory.

It is Lakoff's theory of Idealized Cognitive Models (ICMs) (1987) that underlies both the structure of conventionalized categories as reviewed in this chapter and those non-conventionalized ones investigated in this study. Idealized Cognitive Models are considered as a kind of knowledge base in which our conceptual knowledge is structured and which is "embodied" (Lakoff, 1987: 154), or motivated by our experience. They function to make our conceptual

105

understanding more coherent and thus they facilitate our cognitive functioning, including categorization and reasoning. Cognitive categories are elements in the models and thus linked to each other in such a way that category concepts are organized and structured out of our experience. The taxonomy of object categories is a special kind of idealized cognitive model. The investigation of this kind of categories abounds in experimental psychology and, in particular, in Rosch's famous studies (1973b, 1975). Similarly, non-conventionalized categories, such as the *ad hoc* categories explored in Barsalou (1983) and those investigated in this research, are conceived in ICMs, which grow out of our previous experience and also function to help us manage our interaction with the outside world.

ICMs are created by a particular individual in a particular environment at a particular time. Therefore, ICMs exist in the individual's mind and suggest a way to understand a certain situation, though the way they suggest may not perfectly fit the real world. In a word, ICMs can be regarded as a higher level of conceptual context derived from our prior experience, and subsequently used to explain our new experiences and thus make sense of the world as we encounter it.

As mentioned above, the categories in this study are not conventionalized, but they are more like the *ad hoc* categories investigated in Barsalou's study (1983), which were constructed for a certain purpose, such as the category "ways to escape being killed by the Mafia". The categories under investigation in this study are instantial; they exist in the instances that occur in a particular text and are constructed by the author in his mind and thus require retrieval in the conceptual system of the reader. For example, for the category "*the life of the river*", it is not the surrounding of any common river that concerns the author, but the river in a particular area and at a particular time. In other words, the author created instantial categories which were supposed to be appropriate in the fictional world. The reader

of the text then had to invoke an appropriate context in his/her mind according to his/her understanding of the fictional world. Thus, ICMs have a cognitive status in that they make sense of the world both as created by the author and as comprehended by the reader. To be more specific, the lexical items occurring in the text call up related categories in memory, which in turn call upon corresponding ICMs against which the meaning of the lexical items are defined. Since very few studies investigate this type of category, the current research may provide interesting results with respect to categorization in general and to the explanatory power of ICMs in particular.

In reporting this research, these contextually determined semantic categories would be termed briefly as "semantic categories" because it is well understood that, except experimental research into categories in isolation, every kind of categories occurs in a particular context, and thus is determined by its context.

Chapter 4   The Experiment

4.1   Characteristics of the Research

In the literature on vocabulary learning and teaching, the role of context is prominent and has been widely researched with the aim of showing how it enhances vocabulary learning.   Many applied linguists emphasize the role of the context in which new vocabulary items are embedded, and recommend that learners be advised, or sometimes trained to search for the contextual clues that enable the meanings of these new items to be inferred.   The underlying argument is that an appropriate context can provide sufficient clues for lexical meanings to be inferred, and multiple exposures to context containing the items have been shown to facilitate vocabulary retention (Bensoussan & Laufer, 1984; Nagy *et al.*, 1985; Li, 1988; Mondria & Wit-de Boer, 1991; Coady, 1993; Dubin & Olshtain, 1993; Hulstijn *et al.*, 1996; Prince, 1996).

At the same time, other applied linguists take an opposing position and postulate that isolated vocabulary, i.e. vocabulary items presented without the aid of context, can also increase vocabulary learning, especially when lexical items are arranged in a principled way, as in semantic clustering, or thematic clustering (Channel, 1981; Tinkham, 1997).

The present study focuses on the influence of context on vocabulary learning but from a different perspective than that typically taken.   By re-defining the notion of vocabulary depth knowledge in terms of the ability to recognize contextually determined semantic categories, context is characterized as a cognitive representation constructed by these categories which are invoked by the lexical items encountered by the learners (Ungerer & Schmid, 1996: 46).   The purpose of this research is therefore to investigate the possible relationship between second language learning

and semantic categorization. Very few studies have explored this connection. Cameron (1994) suggested that if the practice of language teaching (L1) could take into account theories of categorization and prototype, it might have some useful implications for L2 teaching. In another empirical study, King (1984) found that the acquisition of sight words was facilitated when these words were at the basic level, as children were able to recall more words overall and more correctly at this level.

However, this researcher knows of no study that investigates the possible role of categorization in second language acquisition, and specifically in L2 vocabulary learning. In this study, successfully L2 vocabulary learning is re-defined in terms of the ability to recognize the structure of contextually determined semantic categories occurring in a text. Because they are so highly constrained by the text, the members of these instantial categories will be very different from those that might be hypothesized in isolation, and therefore this research is therefore likely to represent what occurs in real-life situations, where context will always restrict the concurrent mental representation of categories. Moreover, since the semantic categories under investigation are non-conventionalized and different from the object categories usually investigated in experimental and cognitive psychology, more intriguing results will be obtained and therefore broaden the scope of categorization theory.

4.2 Hypotheses

In this research, two essential properties of category structure are investigated: inclusion and membership. Inclusion refers to the ability of hierarchically higher-level categories to contain lower-level categories, what Rosch terms "the vertical structure of categories" (Rosch, 1978: 30). Membership, on the other hand, is

horizontally presented by the set of more or less prototypical members which constitute the internal structure of a category. Whether language learners can recognize these properties when they encounter them in a text is an empirically intriguing issue which is likely to depend partly on the way the text itself is structured. Basically, lexical knowledge is obtained in two ways, either by referring to real world objects (the typical method favoured in pedagogic materials designed for lower proficiency levels) or by referring to other lexical items presented in the same linguistic context (a method increasingly favoured in pedagogic materials as the proficiency level of the students rises). In terms of vocabulary learning, the former method is associated with breadth and the latter with depth. This experiment explores the role of the linguistic context in facilitating depth knowledge of vocabulary understood in terms of inclusion and membership, and for that reason learners of higher proficiency were chosen as informants for this study.

4.2.1 Inclusion

To test recognition of inclusion properties, the researcher assumes that modifying a given text in a principled way will be likely to affect a learner's ability to infer the hierarchical structure of the categories it contains. It is assumed that an author constructs a text in such a way that the inclusion and membership properties of the vocabulary it contains are maximally transparent to readers so that the text provides a context in which the lexical items in a hierarchy share a dependency which enables them to be absorbed and retrieved more readily. Put plainly, this means that a writer will aim for the maximally efficient way of enabling readers to recover meaning. Specifically, two hypotheses are under investigation. They are:

Hypothesis 1

A text modified so that the inclusion properties of semantic categories are structurally encoded to a greater degree than by the original author will lead to less successful vocabulary learning (= less understanding of inclusion properties) than the author's original text.

Hypothesis 2

A text modified so that the inclusion properties of semantic categories are paratactically encoded to a greater degree than by the original author will lead to less successful vocabulary learning (= less understanding of inclusion properties) than the author's original text.

4.2.1.1 Hypothesis 1

Hypothesis 1

A text modified so that the inclusion properties of semantic categories are structurally encoded to a greater degree than by the original author will lead to less successful vocabulary learning (= less understanding of inclusion properties) than the author's original text.

For the purposes of this research, the term "structural encoding" is used to refer to the way in which a single syntactically structured unit (or sentence) can provide a context within which depth meanings are conveyed. In other words, structural encoding refers to the expression of a message via structure. An example may help to make clear the notion of "a greater degree" of structural encoding. The item *"Derek liked Rachel because she had brown hair"* is regarded as more structured than the original text *"Derek liked Rachel. She had brown hair"* because a structural connective links the two propositions. Sometimes an author chooses to

111

link the higher level items (such as *Derek liked Rachel*) and the included item (such as *She had brown hair*) structurally and sometimes an author chooses to place them side by side paratactically, i.e. in separate structural units, as indeed the author does in this case. The hypothesis states that an author's text which conveys depth meaning by means of paratactic arrangement will lead to more successful vocabulary learning than a modified text in which the same depth meaning is structurally encoded. This is because, as stated at the beginning of 4.2.1, it is assumed that the author knows better.

4.2.1.2   Hypothesis 2

<u>Hypothesis 2</u>

A text modified so that the inclusion properties of semantic categories are paratactically encoded to a greater degree than by the original author will lead to less successful vocabulary learning (= less understanding of inclusion properties) than the author's original text.

"Parataxis" and "hypotaxis" are stylistic terms used to describe sentential structures. Simply put, hypotaxis concerns the structure of subordination, while parataxis does not (Rynell, 1952). Usually, parataxis distinguishes between "asyndetic parataxis" and "syndetic parataxis", with the former referring to a structure without any grammatical connectors and the latter to a structure with coordinating connectors. However, in this experiment, we adopt the definition of Donoghue and Mitchell and take parataxis to refer to the juxtaposition of two or more clauses or sentences without overt connectors (Donoghue & Mitchell, 1992: 169).

For instance, in the structure *it was a long conversation and sometimes she was silent, but sometimes she talked urgently in a low voice*, the higher level item *it was a long conversation* and the included member *sometimes she was silent* were originally linked structurally. The hypothesis states that an author's text which conveys depth meaning by means of structural encoding will lead to more successful vocabulary learning than a modified text in which the same depth meaning is conveyed in paratactic items. Again, this is on the principle that the author knows best how to convey meaning.

4.2.2 Membership

In this research, we are also concerned with the properties of membership, especially two aspects of membership: recognition of membership, (i.e. whether an item is a member of a category) and degree of the membership (i.e. the extent to which an item is a good or less good category member).

4.2.2.1 Hypothesis 3

Hypothesis 3

Successful vocabulary learning (= recognition of the category membership) is more likely to occur where multiple instances of membership at the same level in a semantic category occur alongside co-text which provides the category label, i.e. indicates the level of inclusion of those instances.

Markman designed an experiment in which children were first shown the picture of an object (a standard object) (Markman, 1987). They were then shown two further pictures, one of an object taxonomically related to the standard object, the other of an object thematically related to it, and were required to choose the one

which they thought belonged to the same category as the standard object. Children in the no-word condition were not provided with category labels while those in the novel-word condition were given an unfamiliar label. The latter group were found to choose categorically related pictures significantly more frequently than those in no-word group. The function of the category names used in this study paralleled that in Markman's study in that they were supposed to enhance the recognition of category members. In Hypothesis 3, it is predicted that, when superordinate category names are present, it will be easier for informants to recognize the members of the category and organize the complete category found in the text. Besides, as the theoretical literature claims, multiple exemplars provide a context of comparison which enables informants to infer the structure of the category (Tversky, 1977; Gentner & Namy, 1999), and thus allows for an inductive judgement about the potential properties of other members (Smith & Medin, 1981: 1). Accordingly, there will always be at least two members present in the example tested either with or without category names.

In order to test this hypothesis, the data were selected so as to test the ability to recognize membership in semantic categories containing equal numbers of members either with or without category names specified. For example, in *who had added danger, smuggling and murder to the list of his charms*, the category was explicitly indicated with a category name '*the list of his charms*' for the three members *danger, smuggling* and *murder*, while in *Hey wait a minute!' Oswaldo's large voice bellowed down the phone line. 'Not so fast, OK? Just calm down,'* no category name is specified for the three members, *'Hey wait a minute!', 'Not so fast, OK?',* and *'Just calm down'*.

4.2.2.2 Hypothesis 4

Hypothesis 4

Successful vocabulary learning (= recognition of category membership) is more likely to occur where multiple instances of membership at the same level in a semantic category share common rhetorical properties.

We are also interested in discovering whether the rhetorical properties of the lexical items which are members of a common category affect recognition of category membership to any degree. Within the structuralist tradition, rhetorical "coupling" is regarded as an essential feature of poetic language. "Coupling" refers to an arrangement within which the "natural equivalence" of two or more items, whether phonetic or semantic, is coupled with positional or syntactic equivalence (Levin, 1970: 198). As Jakobson famously stated, *"The poetic function projects the principle of equivalence from the axis of selection into the axis of combination. Equivalence is promoted as the constitutive device of the sequence"* (Jakobson, 1960: 358). Thus, the recurrence of linguistic components, lexical, semantic or phonological, foregrounds these elements (Mukarovsky, 1970: 43). It is assumed in this experiment that coupling helps the learner to notice that a sequence of items belongs to a single category. That is, the coupling of semantic denotation with structural construction will strongly indicate membership of the same category. For example,

It took a bit of time for his words to sink in. It took me time to realise that Malgosia couldn't hear me in there. It took me time to realise that after all my journeying I had finally found her and it wouldn't do me any good at all.

(Trumpet Voluntary, Chapter 8, p. 76)

115

This example illustrates strong semantic and structural coupling of the category members underlined.

4.2.2.3 Hypothesis 5

<u>Hypothesis 5</u>

Successful vocabulary learning (= recognition of degree of category membership) is more likely to occur where multiple instances of membership at the same level in a semantic category have distinct rhetorical properties.

It is widely agreed that the phenomena denoted by lexical items may show different degrees of membership within a category (Rosch, 1973b, 1975; Rosch & Mervis, 1975; Mervis, 1980). This hypothesis tests the ability of learners to differentiate degree of membership in a particular context. According to the structuralist argument, when the message form is foregrounded, the message content is less salient. It is therefore assumed that, when the language is more "poetic", less attention will be paid to the message itself. What Hypothesis 5 therefore tests is whether the ability to recognize the degree of category membership is enhanced when there is less rhetorical coupling, i.e. the lexical items used to denote category members exhibit relatively distinct rhetorical properties.

In brief, Hypothesis 4 predicts that membership of a common category will be more easily recognized when higher degrees of rhetorical coupling occur. However, it is assumed that poetic language helps the reader to notice commonality of category but obscures the degrees of difference in that commonality. Hypothesis 5, therefore, predicts that the degree of prototypicality (the extent to which a member approximates to the prototype of the category) attained by category members will be more readily recognized in cases where the degree of rhetorical coupling is lower.

Both Hypotheses 4 and 5 are designed to investigate the association of rhetorical coupling with category membership.

4.3 Experimental Design

In this section, we will discuss the context in which this research was conducted, the language input that was taken as experimental stimuli, and the informants that participated in this research. However, the measuring instruments will be discussed in a separate section both because of their importance and because of the difficulty in designing suitable instruments to test vocabulary knowledge as defined in this research.

4.3.1 The Context

The research was conducted in a normal classroom setting. The rationale for this is provided first, explaining why such a setting is suitable both for the current research and for research in relation to language learning generally. Second, a discussion of the preparatory work needed to conduct the research will be presented, including the selection of an optimal language course, contact with a language teacher and finally negotiations with the teacher about the implementation of the research. All this was accomplished in such a way as to try to avoid any possible interfering effects on the normal classroom.

4.3.1.1 Classroom-Based Research: general considerations

This research was conducted in an actual language classroom. One of the underlying principles of this research is to find out what actually happens in a real setting, and with the use of regular pedagogic materials. As early as 1983, this requirement was already recognized by Seliger and Long (1983), who discussed the

classroom as the most desirable setting for SLA studies. By emphasizing the context which is the "usual habitat" of language learners (van Lier, 1988: 3), we intend to reveal the reality of language learning and teaching so as to improve our practice in the light of findings, which are very likely to be relevant to language learning and teaching. In short, the language classroom should not be seen as a "black box" any more (Long, 1983: 4), but rather as an accessible and observable environment.

Two aspects of classroom-based research are often criticized. One is the small-scale and short-term nature of the research typically conducted in a normal classroom which inevitably results in data that is small in quantity, leading to doubts as to the reliability of the data and the validity of the research itself. The other is that the data obtained depends on many uncontrollable variables, such as the teaching styles of individual teachers, differences in learners' motivations and attitudes, and even personal learning backgrounds. It is not easy to isolate each of these variables and thus difficult to attribute different outcomes to any single variable or set of variables.

Taking a different point of view, the limited quantity of data typically obtained in classrooms should not be considered a disadvantage because qualitative research does not consider the quantity of data as a crucial element. As argued by van Lier, small quantities of data are still able to reveal significant information about language learning and teaching (van Lier, 1988: 4). In support of this, Long argues that the findings obtained from classroom-based research may provide valuable information with respect to error correction, for example, the form and types of errors, the person who can make the corrections and the best way to correct learners' errors (Long, 1990: 164).

The second limitation is always magnified by those who opt for a laboratory-based research (Hulstijn, 1997b) because difficulties arise in separating individual variables. Preoccupied with isolating each variable, one is liable to neglect the reality of a classroom where no such clear-cut demarcations between interrelating variables exist, and inappropriately apply the results obtained in laboratory-based research where variables can be controlled to classroom contexts where variables cannot be controlled. Thus, avoiding the difficulties originating in the classroom and opting for an artificial setting does not mean that classroom obstacles will be eliminated. In the long run, all findings must be attested in real classrooms if they are to have pedagogic value.

In addition, the generalizability of findings obtained from a single classroom setting may well be questioned. It is always argued that in attempting to make do with whatever is available in a real classroom, the application of such research findings to other situations becomes problematic and any results obtained are thus highly constrained in their wider applicability. Nevertheless, this does not diminish the value of classroom-based research. The understanding of the reality existing in a single classroom can be taken as a single aspect of knowledge concerning classroom practice and together with other such understanding, eventually contribute to the whole picture of language learning and teaching. To acknowledge the shifting and complicated nature of language learning and teaching in all contexts requires us to recognize the importance of the individuality of every single classroom (Larsen-Freeman, 1996: 165).

These problems are often mentioned to justify the practice of laboratory-based research. However, experiments conducted in a laboratory setting also encounter a number of problems. A major weakness of laboratory-based research is the problem of transferability of research findings to classroom settings. Findings

obtained from a well-controlled laboratory will ultimately be challenged by the classroom reality with all its unpredictable variables. In a paper encouraging laboratory-based SLA research because of its superiority of absolute control on potential variables, Hulstijn also recognized the limits of laboratory studies, pointing out that any application of SLA research needs to take into account unpredicted factors (Hulstijn, 1997b: 132). The unpredictable nature of a classroom therefore cannot be neglected.

In addition, laboratory-based research is likely to treat learners as laboratory subjects, rather than as learners in a classroom. This problem exists in recent research on vocabulary learning because learners are required to learn artificial languages which are of no value for their life or academic study, as, for example, in the studies conducted by Tinkham (1997) and Yang & Givon (1997). The usefulness of studying artificial languages by language learners is questionable. Moreover, Tinkham's research imposed a tight control on the time learners were allowed for vocabulary learning, a criticism noted by Waring (1997). The time control imposed for learning vocabulary was neither realistic nor humanistic in terms of language classrooms. These kinds of problems also surfaced in an investigation of the usefulness of bilingualised vs. monolingual and bilingual dictionaries (Laufer & Kimmel, 1997). The researcher presented an incorrect meaning in either the bilingual part (mother tongue translation equivalents) or the monolingual part (target language definition) of a dictionary in order to detect which part was more often used by language learners (1997: 364). Moreover, the presentation of incorrect meanings might be learned by the learner and thus be detrimental to vocabulary development. In such case, an over-concern with the convenience of data elicitation and collection clearly overrides the benefits of learners, and constitutes an ethical problem.

To clarify our position, we are not so much arguing against laboratory-based research as arguing for classroom-based research. A major advantage of laboratory-based research is the control it exerts on the multiple variables inherent in a language classroom, and thus laboratory-based research may contribute to theory in SLA. Nevertheless, classroom-based research should be taken as a precursor of laboratory research as well as an ultimate realization of language learning and teaching theories, because the genuine effects of variables can only be observed in a classroom setting. Moreover, investigation in a natural environment can be seen as more successful in relation to the potentialities of and limitations for language learning and teaching.

In effect, classroom-based research implies a true confrontation with classroom reality, and any results it achieves may serve to reduce the "gap between research and practice" (van Lier, 1988: 15). Since a considerable amount of SLA research is conducted in a laboratory setting, this broadens the disparity between SLA research and practical teaching. This is acknowledged by many prominent SLA researchers and theorists who argue that the application of laboratory-based research is impractical in a language classroom, and furthermore that the relationship between second language acquisition and language teaching is inconclusive and equivocal (Seliger & Long, 1983; Schatcher, 1993; Ellis, 1994; Pica, 1994; Gass, 1995). As a consequence, it is strongly suggested that for SLA to be relevant to language learning and teaching, the "multifaceted nature of the language classroom" (Gass, 1995: 15), or the "totality" of the art (Schatcher, 1993: 185) needs to be accounted for, and that only a concern for teaching practice in the classroom will extend the effects of SLA research and continue to benefit the educational endeavour.

Finally, a word on the role of the teacher in classroom-based research is necessary. As explicitly pointed out by Brumfit & Mitchell, it is the language teacher who knows what is going on in the classroom (Brumfit & Mitchell, 1990:

13). The teacher's cooperation and advice therefore has a significant influence on the success of a research project. To obtain the permission of the language teacher is only an initial but an indispensable step for research experimentation. Taking care of the teacher's feelings before, during and after the experiment is also important. As revealed in recent research projects, the degree of teacher willingness to participate in a research project (Rossiter, 2001), his/her attitudes and conceptions toward the usefulness of research findings relating to classroom practice (Rounds, 1996), and communication between teachers and researchers (Spada, Ranta, & Lightbown, 1996) are significant for the success of research. It is therefore suggested that an effort to establish mutual confidence between teachers and researchers and an attempt to obtain an understanding of the teaching style of the teacher before implementing research is the key to alleviate the tension between teaching and research (Spada *et al.*, 1996: 43).

To summarize the position taken in this research, because many applied linguists and theorists consider classroom settings as the optimal place to investigate second language learning, we conducted the current research in its natural environment. Although practical constraints on the accessibility of language classrooms or on the controllability of complex variables do exist, we are convinced that a language classroom selected for research is able to provide valuable information for second language teaching.

4.3.1.2 The Research Project as An Intervention

In this investigation, the researcher attempted to intervene in the normal classroom practice as little as possible because one intention of the research is to gather data which represents learning in a normal class setting. This is one of the important reasons for choosing graded readers, i.e. regular instructional materials, as

test materials. The Freshman English course in the university from which the experimental sample is drawn requires language teachers to assign whichever readers they consider appropriate for their learners. At the beginning of every semester, teachers receive a list of approved readers and make 3 or 4 choices for their learners. Usually, the length of the readers will be taken as a criterion for choice. Learners are supposed to read a certain length of text each week as an out-of-class assignment, and are expected to complete activities or tasks related to the readers to test their understanding and identify any difficulties they may have with the text. This mode of instruction is used in Taiwan because extensive reading outside class is considered an important means of facilitating second language reading ability and language learning in general (Day & Bamford, 1998: 91).

Bearing in mind the practice as described in the preceding paragraph, the researcher first obtained the agreement of the class teacher to use readers which were considered appropriate both for the purpose of the research and for the level of the learners. The procedures designed to test the hypotheses were also designed as tasks to be done in class-time, thereby making the test procedures as natural as possible in relation to the learning context. In addition, a continuous negotiation was maintained between the class teacher and the researcher before and during the implementation of the experiment.

4.3.1.3 Negotiation With the Class Language Teacher

To avoid giving the informants the impression that they were taking part in an experiment, the researcher had to fully cooperate with the language teacher, including fitting the reading and the test procedures/tasks into her course schedule and making sure that each conformed to her teaching approaches. The Freshman English course consisted of five one-hour classes per week, taught as one class each

day. This meant that the teacher would meet the informants every day for one hour. The teacher allocated one hour per week for the researcher to conduct the experiment, as this conformed to her usual practice for working with readers. Typically, on average 4 readers would be assigned for outside reading each semester (i.e. one per month), and classroom time would be used only for comprehension checks or tests. As the reader contained eleven chapters, two chapters were assigned to be read at a time before class, and the tasks were conducted every other week. Normally, the students came to the classroom and the teacher would check their comprehension either by calling for presentations (in groups or individually) or by requiring written summaries. Both presentations and summaries had to be related to what the students had read in the reader and the topic of the presentation was decided in advance. In the second week, classroom time was allocated for the tasks. Each task took 40-45 minutes.

The researcher was therefore aware that the whole experiment could not occupy too much classroom time, not only because of the limited time available for the researcher, but also because of the status of the instruction materials (i.e. the readers) in the wider course programme, which were looked upon as a kind of supplementary material designed to improve learners' language ability without overt instruction. Other kinds of pedagogic materials, such as coursebooks and grammar practice books, had a preferred status in the overall program, at least in terms of the allocation of class time. Nevertheless, the language teacher always encouraged the informants to be engaged with their reading and emphasized that extensive reading had an important role to play in improving language ability.

4.3.2   Experimental Input

In this section, I will discuss the type of pedagogical materials used in this research, i.e. graded readers.   Specifically, I will draw attention to the advantage of using graded readers for second language extensive reading programmes and their role in enhancing second language learning, and particularly in developing vocabulary knowledge.   I will also discuss related issues such as the characteristics and design of this type of pedagogic material and the debates between those who favour simplified and those who favour authentic texts.

4.3.2.1   Graded Readers

Graded readers have long been used as materials for extensive reading and have proved to be successful in enhancing learners' reading comprehension and in facilitating language acquisition (Hafiz & Tudor, 1989; Tudor & Hafiz, 1989). Their suitability for second language learning is fully spelt out in Hill's review (1997, 2001) and their role in promoting vocabulary learning is empirically accepted (Nation & Wang, 1999).   There are several reasons why graded readers are a useful resource for language teaching.   First, they can promote language learning by preparing language learners for their subsequent encounters with authentic materials outside the classroom.   Second, they strongly motivate learners by providing reasons to continue reading on the assumption that learners are eager to know the ending of the stories (Hill, 1997: 58).   In addition, they are plausible and pleasant materials for an extensive reading programme because they are able to provide "comfortable reading" with nearly 97% coverage of running words (Nation & Wang, 1999: 361) and thus give more confidence to learners reading in a second language (Day & Bamford, 1998: 91).   With regard to the role of readers in promoting second language vocabulary learning, Nation & Wang found that graded readers

enhance vocabulary learning by means of repetition of items at the same level and thus create good conditions for incidental vocabulary learning (Nation & Wang, 1999: 365).

Graded readers are stories written especially for foreign language learners with simple (or, in the case of readers adapted from existing "classic" texts, simplified) lexical items and syntactic structures. They are known as "graded readers" because they are based on a grading scheme which takes into account lexical and syntactic difficulty. Mainstream British publishers typically identify six levels, ranging from elementary to advanced level. The length of individual readers is constrained in accordance with the editing philosophy that specifies the total number of words for each level, the target vocabulary for each level and an ideal reading programme. Normally, each reader contains repeated lexical items. The underlying assumption is that extensive reading will facilitate language acquisition by means of multiple exposures to new lexical items presented in a comprehensible and engaging context.

Although graded readers have achieved popularity recently, a controversy still remains. The central issue underlying the use of graded readers is whether simplified texts can take the place of authentic materials. These simplified versions have been around for a long time and the resurgence of interest in readers has been stimulated in part by the relatively recent appearance of "original" fiction written specifically for language learners. Some applied linguists have criticized simplified texts as not being natural discourse in that they are organized according to pedagogic principles and tend to avoid the redundancy found in everyday discourse (Nuttall, 1996: 178); others have not found that simplified text could facilitate language learning (Leow, 1993). During the process of simplification, priority is placed on a reduction in the range and complexity of the language code; something is felt to be lost in comparison with native speaker "genuine" discourse. Thus, it is suggested

that too much emphasis on a reduced code may adversely affect the content and the naturalness of the discourse. However, authenticity is not incompatible with simplification since it is the quality of writing that constitutes the essence of authenticity (Day & Bamford, 1998: 61). That is to say, simplified texts can also possess the quality of authenticity if written with skill (Alderson & Urquhart, 1984: 198). In distinguishing between "a simplified version" and "a simple account", Widdowson indicates the possibility of "a simple account" being "genuine discourse" (Widdowson, 1978: 89). Thus, referring to an original text written in simple English, Hill renames Widdowson's "simple account" a "simple original" (Hill & Thomas, 1988: 44).

Moreover, for language teachers, the pedagogical merit of simplified texts overrides the importance of authenticity. Whilst authentic materials often appear so difficult as to distract learners' attention from the intended message and cause them to focus on problematic language items, simplified texts can provide an alternative suitable for language learners with limited language proficiency (Nuttall, 1982; Williams, 1983; Wodinsky & Nation, 1988). The breakdown of communication between writer and reader caused by language difficulty may deprive authentic materials of their obvious value, whilst a fluent reading may on the other hand be enabled by the use of a simplified text. Breen points out that the quality of authentic texts has to be related to the language learners' ability to understand the texts (1985: 61). Learners may "authenticate" (*ibid.* 64) a text during the process of interpretation. As Davies concludes, "it is not that a text is understood because it is authentic but that it is authentic because it is understood" (Davies, 1984: 192). Another benefit of using simplified texts at an early stage of language learning is that they can provide with an easy bridge to real reading so that when language learners

read with comfort, they are more strongly motivated to read more (Davies, 1984; Hill, 1997).

Whatever the merits of the various positions in the debate, many publishers are involved in the production of graded readers, and some series have achieved considerable popularity.   In this investigation, it is assumed that, as the publishers intend, graded readers are able to provide a self-contained context for vocabulary learning in relation to linguistic categorization.

4.3.2.2   The Selection of Graded Readers

As Read argues, there is no way of knowing how much information is necessary for contextualized vocabulary learning (2000: 187).   This is especially true for this research in which vocabulary knowledge is equated with the ability to recognize categorical structure in a text rather than to learn new items in isolation. In addition, the selection of texts will necessarily be based on what is available in practice.   In this experiment, the researcher had the opportunity to preview a large number of graded readers from which to choose a text suitable for experimental purposes and appropriate to the informants' proficiency level and perceived interests.

The present experiment was designed to test more advanced vocabulary knowledge, i.e. depth of vocabulary knowledge measured in terms of the ability to recognize the inclusion and membership properties of semantic categories.   We limited our choice to readers at levels 5 and 6, which were considered suitable for upper-intermediate or high-level learners.   Level 5 readers were limited to 2800 headwords (the base form of a word), while level 6 was limited to about 3800 headwords.   Each level included several readers covering various genres, such as romance, thriller, murder and mystery.   Another group of readers, simplified

readers based on or related to films, were eliminated from our consideration because they were unavailable for modification for copyright reasons.

The final selection was made from those readers with rich categorical structure and themes of interest to the informants. Jeremy Harmer's *"Trumpet Voluntary"*, a level 6 reader, was eventually chosen as it met the criteria better than the other texts considered.

4.3.3 Informants

In this section, I will describe the factors behind the selection of the group of second language learners as informants for this research, the procedures of experimental grouping for testing different hypotheses, and the language learning background of the learners. It is worth noting that prior to and during the research, informants were not told anything about the experiments they were to participate in to avoid any unpredictable motivational effects. Indeed, they were not aware until the data had been collected that they were not involved in regular instruction.

4.3.3.1 Sampling

In order to test the five hypotheses listed in section 4.2, a university setting was selected because the experiment required informants of higher language proficiency in order to test depth of vocabulary knowledge as revealed by the informants' ability to recognize the inclusion and membership properties of semantic categories. An experimental class was selected on the basis of high language proficiency, as measured by the university's English placement test.

The advantage of this method of sampling is that the informants constituted groups who were taught and who learned together in a real classroom environment, thus enabling the experiment to be conducted in a natural rather than a "laboratory"

setting. In addition, the class was not informed in advance that their regular learning was also being studied for research purposes so as to eliminate any psychological effects such as reluctance to participate in an experiment. Detailed personal information about the students was obtained via questionnaires after the experiment was conducted (Appendix A), at which point the informants were asked to indicate their willingness for the data they had provided to be included in the experimental data (Appendix B).

In order to test the effects of modification on the text (Hypotheses 1 and 2), the class was divided into two sub-groups, A & B. For Hypothesis 1, sub-group A read the author's original text and sub-group B read a modified version. For Hypothesis 2, sub-group B read the author's original text and sub-group A read a modified version. Neither of the sub-groups was told that they were going to read different versions of the text. The text was printed out as handouts and distributed among the informants chapter by chapter to be read as an assignment task. Thus, the informants were required to read and study the text out of class in preparation for the class in which the language teacher would check their understanding of the text, either by tasks such as those designed in this research or by some classroom activities.

4.3.3.2 Language Learning Background

At the time of the experiment, these informants had been studying English for at least six years in high school. The exposure time varied between six and nine hours per week, depending on the different academic or vocational backgrounds of the students. After six years of learning in high school, the students were expected to be able to read and write simple English at proficiency levels ranging from basic to intermediate.

However, grammar-translation is still the prevailing instructional method in high schools in Taiwan. That is to say, most students did not have much chance to practise speaking in class, and thus they were much less proficient in speaking than in other skills such as reading and writing and so tended to be highly anxious about speaking in public. A major difference between the high school and college level English courses is that most colleges and universities emphasize speaking ability in Freshman English Courses in order to remedy the perceived deficiency.

One point deserves mentioning is that this group of informants have acquired a linguistic categorization device along with their mother tongue, Mandarin. Like other languages in south-east Asia, the classifier system in Mandarin deals with classification of objects on the basis of perceptual features and is very different from Western languages. This system may or may not affect the categorization behaviours of this group of informants, and therefore a brief introduction of Mandarin classifiers will be provided in Appendix C.

4.4 Instrument of Measurement

In this section, I will describe how we designed different tasks for each hypothesis in relation to the target items that have been identified earlier. Although we could not find any workable instruments to test the type of vocabulary knowledge specially defined in this research, we looked to the vocabulary testing literature for inspiration about the format of the tasks we needed to design. Each task/test will be discussed in turn and typical items identified in the reader presented as illustrations. We also explain why these task types and test items selected are an appropriate means of testing the hypotheses. Moreover, we explain how each task was constructed as a specific means of testing each hypothesis. Finally, the contribution of each task to the language classroom will be discussed

4.4.1 Vocabulary Assessment in Previous Research

In this section, I will first discuss the instruments used for measuring vocabulary knowledge as revealed in the literature with a view to exploring their suitability for the present research, especially bearing in mind that vocabulary knowledge in this investigation is defined in terms of the ability to recognize the inclusion and membership properties of semantic categories in a particular text. The existing instruments are widely used in empirical research and recognized as reliable ways of measuring vocabulary size (breadth), vocabulary depth and productive use. The discussion that follows will be limited to a discussion of the five most influential contemporary tests. (For a wider review of test types, see chapter 2, section 2.4). They will be classified into three groups, two that measure vocabulary size, two that measure vocabulary depth, and one that measures language production.

1) Vocabulary size: the Vocabulary Levels Test (Nation, 1990), the Eurocentres Vocabulary Size (Meara & Jones, 1990)

2) Vocabulary depth: word association test (Read, 1993; Schmitt & Meara, 1997), and Vocabulary Knowledge Scale (Wesche & Paribakht, 1996)

3) Productive Vocabulary Knowledge: the Lexical Frequency Profile (Laufer & Nation, 1995)

The instruments in group 1 are generally not suitable for the present research in that the vocabulary knowledge as defined in this investigation is related more to vocabulary depth than to breadth. The Vocabulary Levels Test consists of five parts, corresponding to five levels of word frequency 2000, 3000, 5000, 10000 and the Academic Word List, as proposed by Nation (1990). Word-definition matching is the format used to construct the test. That is, informants are required to match

test words to possible definitions. Each test consists of six test words and three possible definitions. There are 36 test words and 18 definitions in total for each level. According to Nation, this design reduces the chances of providing correct answers by guessing (Nation, 1990: 162). This type of level-division test is used mainly to differentiate levels of English proficiency on the basis of knowledge of frequently occurring words. It is therefore a measurement of the breath of vocabulary knowledge.

The Eurocentres Vocabulary Size Test, on the other hand, uses a yes/no format to test word-recognition. It is actually a checklist covering a large number of words, and includes non-words. It includes items at the 1K (i.e. 1000), 2K, 3K, 4K, 5K levels and in the Academic List. As words appear on the computer screen, an informant is required to answer 'Yes' or 'No' by pressing an appropriate key. The role of non-words is to enable the scorer to adjust the possible score if informants overestimate their vocabulary knowledge, as indicated when they signal recognition of a non-word. Nevertheless, both Nation's Vocabulary Levels Test and the Eurocentres Vocabulary Size Test only measure word-recognition ability and are therefore not appropriate for this research into knowledge of inclusion and membership. Besides, as explained in chapter 3, the lexical items covered by the five hypotheses in this research are not "words" as they are conventionally thought of, but rather units of multiple words.

In the second group are instruments used to measure vocabulary depth. Word association tests have been long established as a way of investigating associative behaviour in response to stimulus words. Researchers have used the tests to elicit various types of responses to both L1 and L2 stimulus, such as coordinates (words included in the same category), similars (words of nearly the same meanings) or contrasts (words of opposite meanings) (Riegel, Ramsey, & Riegel, 1967: 537-538).

More recently, efforts have been made to investigate the relation between language proficiency and word association behaviour, as in the study of Kruse, Pankhurst, and Sharwood Smith (1987), and the degree of nativeness obtained by non-native speakers with respect to word knowledge (Meara, 1984; Read, 1993). The latter study involves a comparison of responses between native speakers and non-native speakers. The traditional procedure is to give subjects a stimulus word and ask them to produce the first words that come to their mind. These responses are later compared with norm responses obtained from native speakers. Nowadays, more than one response is sometimes required, as in Schmitt & Meara (1997), whose study required the subjects to list three associated words; or sometimes subjects are required to choose semantically associated items from a set of given items as in Read (1993). As a result, various kinds of semantic relations between the stimulus word and response words can be tested in this kind of vocabulary assessment. However, this test type measures word depth in a free context and thus cannot be used in this research, where context has an important role in defining target items as organized within a hierarchical structure. Accordingly, word association tests are not ideal for this research.

Another kind of instrument, Wesche & Paribakht's Vocabulary Knowledge Scale (1996) measures the depth of the vocabulary knowledge of test-takers via self-report in a five-level format, beginning with superficial recognition of test items (level 1) and progressing to appropriate and grammatical use (level 5). The unsuitability of VKS for the present investigation stems from the different definitions of depth of vocabulary knowledge used by Wesche & Paribakht and the present researcher. In Wesche & Paribakht's work, vocabulary depth is defined in terms of appropriate and grammatical use of test items (level 5), but in this investigation, depth is defined in terms of the ability to recognize the inclusion and

membership properties of contextually determined semantic categories. Nevertheless, the format of different levels for measuring degrees of vocabulary knowledge was found worth adopting, as will be explained subsequently.

The first two groups of tests basically measure the recognition of words to a more or less degree, while the last test considered here measures productive knowledge. Laufer and Nation's Lexical Frequency Profile (LFP) mainly measures lexical richness in free composition (1995). In this test the researchers differentiate the proficiency levels of learners by setting up profiles based on calculating the percentages of words occurring at the frequency levels of the first 1000, the second 1000, the University Word List (UWL), and not-in-the-lists (i.e. much less frequent words) out of the total number of tokens used. Accordingly, the researchers can differentiate the proficiency levels of language learners. Since this test is basically concerned with the productive knowledge of vocabulary, it is not suitable for the present research, which is concerned more with recognition of semantic categories.

Even though the present research is not concerned with vocabulary depth or breadth in the usual sense, the level approach to vocabulary depth adopted in the VKS and, despite being designed to measure breadth, Nation's Vocabulary Levels Test both provided formats that were useful starting points for test design in the present investigation, as described below. As has been explained, no single existing test was found to be entirely suitable for this investigation because of the original conception of contextually determined semantic categorization which was adopted by the researcher. It was therefore necessary to design new instruments of measurement appropriate to the experimental design, although the researcher was fully aware of the difficulty of constructing a vocabulary test, as noted by Wesche & Paribakht (1996). The instruments of measurement used in this research are therefore largely original and designed to meet the specific needs of categorization

within a cognitive linguistic view of lexical knowledge. Inevitably, therefore, they are more open to evaluation as instruments of measurement than might otherwise be the case.

In designing these original instruments, we tried to keep in mind Read's suggestions that four aspects should be taken into consideration while designing a vocabulary test: 1) format either simple or complex; 2) answer or response either verifiable or self-report; 3) vocabulary knowledge tested-either breadth or depth of; 4) target items tested-either in context or in isolation (Read, 1993: 355-57). Since it is stated clearly from the start that this research is an investigation of vocabulary depth and context will be a constraining factor in an extended text, the answers to criteria 3) and 4) are evident. In addition, this research adopted a simple format and verifiable responses in order to facilitate the collection and analysis of data. A simple format can be more transparent for the informants who are therefore more likely to understand the task requirements, especially when time limits inherent to the classroom setting are imposed on the research as in this case. Moreover, since there are many doubts about the reliability of self-reported data, the present researcher tried to build into these instruments of measurement a degree of objective verifiability of the data provided, although fundamentally relying on self-report. The degree of verifiability was obtained by asking learners at the outset of one particular test whether they could understand the superordinate category term in isolation. As a result of the information provided in response to this question, the researcher was able to compare the subsequent data provided by informants who claimed to understand and who stated that they did not understand the item in question.

4.4.2 Target Items

In this section, we will explain only how target items were selected for each hypothesis from sections 4.4.2.1 to 4.4.2.4. It involved first identifying potential items from the reader and then selecting suitable ones that displayed more transparent category structure. Detailed procedures and special considerations for a particular task will be provided. However, the construction of each task will be described in section 4.3.3. That is, the design and implementation of each task in relation to hypothesis testing will be discussed from sections 4.4.3.1 to 4.4.3.4.

Supported by past research showing that hierarchical structure also exists in daily life events and scenes (Rifkin, 1985; Morris & Murphy, 1990; Tversky & Hemenway, 1983) within a superordinate-basic-subordinate structure, the researcher identified instances of superordinate/basic and basic/subordinate contrasts in the text of the graded reader as potential test items for Hypotheses 1, 2 and 3, and instances of two or more members of a common category as potential test items for Hypotheses 4 and 5.

To test the five hypotheses systematically, four chapters were chosen from the text which contained 11 chapters altogether. It was decided that Hypothesis 1 would be tested by data drawn from chapter 3, Hypothesis 2 by data drawn from chapter 4, Hypothesis 3 by data drawn from chapter 7, and Hypotheses 4 and 5 by data drawn from chapter 8. Chapters 1 and 2 were excluded on the grounds that until the informants had sufficient familiarity with the broad context invoked by the reader, it was inappropriate to test for inclusion and membership, given their dependence on both wide and narrow contexts. A further consideration was the desirability of spreading the tests/activities throughout the teaching period. These considerations apart, the choice of chapters was not random but depended upon whether there were sufficient target items present to enable the hypotheses to be

tested. For example, chapters 3 and 4 were chosen to test inclusion because they contained a sufficient number of items structured with inclusion explicitly indicated; chapter 7 provided a contrast between items which were presented with and without inclusion indicated through the presence of a superordinate category name and was therefore selected to test Hypothesis 3. In chapter 8, there were a large number of instances where two or more members of a category were listed and these instances exhibited a range of degrees of rhetorical coupling.

4.4.2.1 Target Items for Task 1

Task 1 was designed to test Hypotheses 1 and 2, which were used to test the ability to recognize the inclusion properties of the semantic categories established in the text.

Hypothesis 1

A text modified so that the inclusion properties of semantic categories are structurally encoded to a greater degree than by the original author will lead to less successful vocabulary learning (= less understanding of inclusion properties) than the author's original text.

Hypothesis 2

A text modified so that the inclusion properties of semantic categories are paratactically encoded to a greater degree than by the original author will lead to less successful vocabulary learning (= less understanding of inclusion properties) than the author's original text.

Hypothesis 1 states that an author's text which conveys depth meaning by means of paratactic arrangement will lead to more successful vocabulary learning than a modified text in which the same depth meaning is structurally encoded. On the other hand, Hypothesis 2 states that an author's text which conveys depth meaning by means of structural encoding will lead to more successful vocabulary learning than a modified text in which the same depth meaning is conveyed in paratactic items. Therefore, we first identified cases where the basic level items that were considered members of the category and the superordinate category name occurred in paratactic structures (Appendix D)[1], and then recreated the text so that these items were more structurally encoded (Appendix E). To construct structural test items, we added coordination or subordination connectors as appropriate to link formerly separate sentences. For example,

That night we sat and talked about what we hoped for the future. I told her I wanted to make enough money as a musician to have a nice house, travel a bit, that kind of thing.

(Trumpet Voluntary, Chapter 3, p. 23)

was re-written as:

That night we sat and talked about what we hoped for the future, so I told her I wanted to make enough money as a musician to have a nice house, travel a bit, that kind of thing.

In contrast, in chapter 4, we identified cases where the basic level items that were considered members of the category and the superordinate category name occurred in a single structural unit and modified these units so that the superordinate category name and the category members occurred in paratactic structures. To

---

[1] Appendix C contains the author's text, which is given in full, so as to enable the reader to judge the

construct paratactic test items, we typically eliminated coordination or subordination connectors between clauses and juxtaposed these clauses as separate sentences. As an example, the sentences:

...but it was a long conversation and sometimes she was silent, but sometimes she talked urgently in a low voice.

(Trumpet Voluntary, Chapter 4, p. 34)

was modified to yield:

...but it was a long conversation. Sometimes she was silent, but sometimes she talked urgently in a low voice.

In total, six items from chapter 3 and another six from chapter 4 were selected to test Hypotheses 1 and 2 respectively. These items were arranged randomly so that they did not follow the order in which they were encountered in the text, except that those in chapter 3 testing Hypothesis 1 preceded those in chapter 4 testing Hypothesis 2.

4.4.2.2   Target Items for Task 2

Task 2 was designed to test Hypothesis 3, which was used to test the ability to recognize the membership properties of the semantic categories with/without the presence of category names.

Hypothesis 3

Successful vocabulary learning (= recognition of the category membership) is

more likely to occur where multiple instances of membership at the same level

---

appropriateness of the examples chosen by the researcher.

in a semantic category occur alongside co-text which provides the category label, i.e. indicates the level of inclusion of those instances.

This hypothesis states that when superordinate category names are present, it will be easier for informants to recognize the members of the category and organize the complete category found in the text. Therefore, it was also necessary to identify sequences where the category names were present and sequences where they were not. Many instances of each type, i.e. with or without an indication of the level of inclusion, were identified. It was predicted that, when superordinate category names were present, it would be easier to recognize the members of a category in the text. However, Gentner and Namy demonstrated that the number of exemplars made a great difference in category abstraction (Gentner & Namy, 1999). Thus, it was necessary to identify sequences of two or more members of a single category as test items in chapter 7. Care was taken to make sure that the number of test items containing a certain number of members with a category name was matched by an equal number of members without a category name. Thus, the number of items containing 2, 3 or 4 members with a category name was matched by the same number of containing 2, 3, or 4 members without a category name. Thus, two-item (i.e. 2-member) examples with category name indicated, such as

That's why I probably behave a bit stupidly in dramatic situations, whether at Rosemary's front door or on a Brazilian hillside. (2 items, with category name present)

(Trumpet Voluntary, Chapter 7, p. 68-9)

were matched by two-item examples without category name, such as

'She means that your Malgosia isn't here anymore, that's what,' Oswaldo said bluntly. *'She's gone.'* (2 items, without category name)

<p align="right">*(Trumpet Voluntary, Chapter 7, p. 72)*</p>

In total, eighteen items were selected from a set of identified potential test items to test Hypothesis 3. Instances containing 2 items appeared first, then those containing 3 items, followed by those containing 4 items. Within each set, the examples were randomly ordered with respect to those where category names were specified. The numbers of different items are listed below:

Table 4.1: Numbers of different test items in Task 2.

Types	2 members with category name	2 members without category name	3 members with category name	3 members without category name	4 members with category name	4 members without category name
Numbers of items	4	4	2	2	3	3

4.4.2.3 Target Items for Task 3

Task 3 was designed to test Hypothesis 4, which investigated the extent to which common rhetorical properties affected the ability to recognize category membership.

Hypothesis 4

Successful vocabulary learning (= recognition of category membership) is more likely to occur where multiple instances of membership at the same level in a semantic category share common rhetorical properties.

This hypothesis states that the more common the rhetorical properties shared by category members, the more likely category members are to be recognized as such. That is, a higher degree of coupling helps the learner to notice that a sequence of items belongs to a single category. In the initial stage, a total number of 23 potential test items was identified in chapter 8, showing varying degrees of rhetorical coupling. In order to determine the degree of rhetorical coupling of each item, a sample of native speakers was asked to act as raters. The raters were 2nd and 3rd year undergraduates studying in the Schools of Linguistics or Education at the University of Durham. Because the raters were drawn from students registered on an English Language Teaching module, they were expected to have a professional interest in the research with which they were assisting. They were each given a letter explaining the purpose of the experiment and a short training section instructing them how to rate the test items based on their intuition (Appendix F). A 4 point scale was provided ranging from "0" to "3" where "0" meant no detectable coupling and "3" indicated strong parallelism. After the training section, the informants were asked to do the ratings out of class. 12 students returned the rating sheets and their scores were aggregated and divided by the number of raters. The results of average ratings for all test items are listed in Appendix G.

In this section, we discuss the items selected in Task 3 to test Hypothesis 4. Because Hypothesis 4 is concerned only with the possible association of recognition of membership and degree of rhetorical coupling, only items where no superordinate category name was present were selected so as to eliminate the possibility that the presence or absence of a category name might have an influence on the data elicited from the informants. The following example:

both typical symptoms of someone who had been poisoned by <u>some form of nerve gas,</u> <u>some kind of chemical agent</u>. (average rating: 2.6)

(Trumpet Voluntary, Chapter 8, p. 79)

showed a higher degree of coupling signalled by reoccurrence of the same structure and to some degree of the semantic denotation. Whether the properties of category membership present in this item were more readily recognized than in an example with a lower degree of coupling was under investigation. Virtually, no perceptible coupling was present in:

Her hair was cut short. She wore heavy black glasses. <u>She watched me as I struggled</u> <u>in with my suitcase. She didn't look welcoming at all</u>. (average rating: 0.4)

(Trumpet Voluntary, Chapter 8, p. 74)

It was hypothesized that it would be more difficult for the informants to recognize the members of the category in an example such as this than in an example where the coupling rating was higher for the task. All the items were randomly arranged so that they did not follow the order in which they were encountered in the chapter.

4.4.2.4   Target Items for Task 4

Task 4 was designed to test Hypothesis 5, which was used to test the association of the extent to which distinct rhetorical properties present among category members might affect the informants' ability to differentiate degree of category membership.

<u>Hypothesis 5</u>

Successful vocabulary learning (= recognition of degree of category membership) is more likely to occur where multiple instances of membership at the same level in a semantic category have distinct rhetorical properties.

This hypothesis states that, when the lexical items used to denote category members exhibit relatively distinct rhetorical properties, the informants will be more likely to differentiate degree of category membership. That is, a lower degree of rhetorical coupling among category members would lead to greater differentiation of degree of membership and a higher degree of rhetorical coupling would result in less awareness of degree of membership. The test items were a subset of data drawn from chapter 8 as previously identified. As stated above, items with category names were used in Task 4, since the informants were specifically required to distinguish degree of membership within an identified category. Two candidate items from the set originally identified and rated were eliminated because the researcher was uncertain about their category structure. In total, 11 items were used. The following example in which two causes of death were specified had a very low average rating of 0.5.

what Malgosia's father told me was that my wife had died, in the end, <u>bleeding unstoppably</u>, <u>completely unable to breathe</u>. (average rating: 0.5)

(Trumpet Voluntary, Chapter 8, p. 79)

It was expected that the informants would assign to the two members, *bleeding unstoppably* and *completely unable to breathe*, degrees of membership across more levels, such as between the levels of **a very good** and **not a good,** because of this lack of common rhetorical properties. In contrast, the following example in which two undesirable characteristics are listed was expected to cause less awareness of

discrimination among category members because two members, *her bright-eyed intelligence* and *her deep dark passion*, possessed some common rhetorical properties as shown in the average rating of 1.9.

It made me realise where Malgosia got <u>her bright-eyed intelligence</u> and <u>her deep dark passion</u>. *(average rating: 1.9)*

(Trumpet Voluntary, Chapter 8, p. 73)

The researcher was aware that the informants might draw on their cultural background knowledge in judging the degree of membership for each category member. In addition, in each item there would be an inherent degree of prototypicality for each category member, with the possibility that in some cases it would not be appropriate to differentiate degree of membership. The researcher was fully aware of these possibilities and therefore included as large a number of test items as possible and scored the results according to the ability of the informants to discriminate difference irrespective of which members were assumed to be better or less good members of the category and irrespective of the degree of difference asserted.

4.4.3 Classroom Tasks and Hypothesis Testing

Four classroom tasks were designed to test vocabulary knowledge. To present a clear picture of the whole design of the experiment and the relationship of the various tasks to each hypothesis, we list below the sequence of lessons that took place in the classroom (column 1), the tasks these lessons contained (column 2), the hypotheses that the tasks were used to test (column 3), and the chapter in the graded reader of the task materials were based (column 4), and the corresponding experiments as they will be referred to in chapters 5 and 6 (column 5). The reader

may find the table useful subsequently in understanding the relationship between numbered lessons and tasks in the sequence as well as numbered hypotheses and experiments.

Table 4.2:   Numerical coding of Tasks, Hypotheses and Experiments.

Lesson	Task	Hypothesis	Data from	Experiment
1	1 (Pt 1)		Ch. 3 Ch. 4	1
2	1 (Pt 2)	1 2	Ch. 3 Ch. 4	2a 2b
3	2	3	Ch. 7	3
4	3	4	Ch. 8	4
5	4	5	Ch. 8	5

Since the text was not "taught" in the class, the tasks were considered to be an alternative to classroom instruction and intended as an "input enhancer" (Gass, 1997: 154) with the aim of drawing the learners' attention to the target input. Tasks in this experiment were used to raise learners' awareness of the lexical knowledge they had obtained from the input. They functioned to bring language awareness as measured by the tasks to a conscious level, and encouraged the learners/informants to reflect on the patterns already established during their initial reading of the text. These tasks were therefore indistinguishable from the type of exercises practised in regular classrooms in the institution concerned. As noted earlier, this was one of the characteristics of this experiment which distinguished it from the laboratory type experiments more commonly used to assess vocabulary knowledge.

Task 1 was designed to test the ability of informants to recognize inclusion properties either by virtue of structural or of paratactic encoding of the category name and its members. Tasks 2, 3 and 4, on the other hand, were designed to test

147

the ability of informants to recognize category membership properties. Task 2 was designed to investigate whether the presence of superordinate category names made any difference to recognition of category membership, and Tasks 3 and 4 were designed to measure whether rhetorical coupling among category members could enhance recognition of category membership whilst impeding further differentiation of category membership.

4.4.3.1 Task 1

As stated, Task 1 was designed to test Hypotheses 1 & 2 and specifically to discover whether an association could be established between structural and paratactic principles of text construction and the understanding of inclusion. In order to test these hypotheses, the level approach to Vocabulary Knowledge Scale pioneered by Wesche & Paribakht (1996) was adopted as a starting point for the design of the instrument of measurement. Wesche & Paribakht's original VKS consisted of five levels, each representing a different degree of vocabulary knowledge depth. At the first level, it tested the superficial recognition of lexical items. Similarly, Task 1 (see Appendix H) first required the informants to indicate whether they could understand the literal meanings of the target items which were treated as category names by the author of the reader. If informants reported that they could understand the item, they went on to a second question requiring them to list any two examples that they considered could be included in the category. However, if they stated that they did not understand the target item, they were instructed to skip the second question and go on to next test item. This was the Part 1 of the task. The example shows how the tasks were presented to the informants.

Question 1   *a messy life*

Tick either (a) or (b) below
a) I don't understand this phrase
b) I think I understand this phrase

If you have ticked (a), go straight to the next question
If you have ticked (b), you should also answer (c)

c) The kinds of thing I think of when I think of a messy life include
Example 1 ....
Example 2 ....

The reason for setting the task up in this way was to ensure that the informants could indeed understand the target items which had been extracted from their context and furthermore that self-report could be cross-referenced with Part 2 to determine the reliability of the answers given to the questions in Part 2, which would be described below. This strategy therefore enabled the researcher to determine whether items which the informants considered they did not understand when presented out of context (Part 1) yielded different results when presented in context (Part 2) from the results provided by the informants who claimed to understand the item out of context.

Before Part 2 of Task 1 was administered, a training section was provided to explain the notion of inclusion (see Appendix I). Specimen examples were drawn from taxonomic categories and event categories in daily life such as *The animals we saw yesterday were elephants, lions, hippos and giraffes*. The instruction and examples were written in Chinese and English to ensure that the informants could grasp the idea of inclusion.

In Part 2, all the test items appeared again followed by their exemplars taken directly from the text. For each example, four choices were provided and the informants were asked to indicate whether they thought they had read the example or not and whether, according to their opinion, it could be included in the category

name given above.   The informants were allowed to choose only one answer for each example.   The following example is provided as an illustration:

Question 1   *Derek's messy life*

Examples:
Example 1 he was playing badly   A B C D
Example 2 he didn't have many friends   A B C D
Example 3 the person he thought he loved had disappeared from his life   A B C D

Key:
A means 'I think I read this and it could be included in Derek's messy life'
B means 'I think I read this and it could not be included in Derek's messy life'
C means 'I don't think I read this and it could be included in Derek's messy life'
D means 'I don't think I read this and it could not be included in Derek's messy life'

Recall that the informants had read one of two texts when they attempted these task items, either the original or the modified one.   These four choices made it possible to test the ability of the informants to recognize inclusion and also to evaluate the contribution of reading effects on the ability to recognize inclusion. Part 1 and Part 2 were administered in separate sessions to prevent informants drawing on the examples given in Part 2 as answers to Part 1, thus making Part 1 invalid.   For example, in Part 1 for the question of *a messy life*, the informants were prevented from listing any examples given in Part 2 such as *he was playing badly* or *he didn't have many friends*.

4.4.3.2   Task 2

Task 2 (Appendix J) was designed to test Hypothesis 3, i.e. to test the informants' ability to recognize the membership properties of semantic categories. It required the informants to identify all the members belonging to the same categories in the discourse context in which they occurred.   The informants were provided with the entire context in which the members to be identified were

embedded in the original text and were asked to indicate the members of the common category by bracketing them. The number of members in a category was explicitly indicated in each question. For instance,

Question 3 (2 items)
It was quiet too. Occasionally a snatch of conversation from one of the houses came to us on the light wind, or the sound of a car in the distance.

As this example indicates, there could also be some discussion as to just which items might be bracketed. For example, would *a snatch of conversation* or *a snatch of conversation from one of the houses* do as well as *a snatch of conversation from one of the houses came to us on the light wind*? In fact, all three of these suggested bracketings would have been accepted by the researcher because all three show that the informant had clearly recognized one of the two common category members. This principle was adhered to throughout, i.e. wherever an informant had clearly recognized a category member, the answer was scored as correct even though less than the full phrase might have been bracketted.

A short training session (both in English and in Chinese) was provided which once again explained the notion of category membership with examples drawn from taxonomic categories and categories in daily life events. The teacher demonstrated how to bracket each member as shown in the example below:

Specimen Question 2
My favourite fruits are bananas and mangos (2 items)
（我最喜歡的水果是香蕉和蘋果。）

Correct Answer
My favourite fruits are [bananas] and [mangos]
（答案：我最喜歡的水果是【香蕉】和【蘋果】。）

4.4.3.3 Task 3

Task 3 (Appendix K) was designed to investigate any possible association between the ability to recognize category members and the rhetorical properties of the text that conveyed them, thus testing Hypothesis 4.   As in task 2, the informants were also required to make a judgment on which items were the members of a common category.   However, Task 3 differed from Task 2 was that informants were provided with three possible sets of category members presented in a multiple-choice format.   This format therefore invited the informants to judge membership between three alternatives only one of which contained only members of a common category. Task 3 therefore tested only recognition of membership so that the results would be correlated with the NS ratings described above (see p. 143).   An example is given below:

Question 5
a) Anja started up a low terrible moaning which seemed to go on and on. I wished she would just stop. <u>Even when her father made me go with him into the kitchen we could still hear it, the inhuman noise of someone who was going to feel guilty for ever.</u>
b) Anja started up <u>a low terrible moaning which seemed to go on and on</u>. I wished she would just stop. <u>Even when her father made me go with him into the kitchen we could still hear it,</u> the inhuman noise of someone who was going to feel guilty for ever.
c) Anja started up <u>a low terrible moaning which seemed to go on and on</u>. I wished she would just stop. Even when her father made me go with him into the kitchen we could still hear it, <u>the inhuman noise of someone who was going to feel guilty for ever.</u>

Because only the underlined items in (c) belong to a common category, (c) is the correct answer and (a) and (b) are incorrect.

4.4.3.4   Task 4

Task 4 (Appendix L) was designed to investigate the possible association between relative absence of common rhetorical properties amongst sets of category members and the ability to differentiate degrees of membership within a category. In Task 4, each test item consisted of a category containing both category name and a set of members.   The informants were to rate the prototypicality of each member of the category.   The format of this test followed that of Nation's Vocabulary Levels Test. However, whereas Nation's test required informants to match words and definitions, this test required them to match items and descriptions of degree of membership.   In the Vocabulary Levels Test, six test words were provided together with three matching definitions.   In this task, different numbers of examples were supposed to be assigned a relative degree of membership out of four possible options. Thus, this test allowed the researcher to determine whether the informants were able to differentiate each member in relation to their degree of membership of a category. In the following example, the informants were therefore asked to circle for each of the two given category members, the description that best accorded with its degree of membership as it appeared to the informant:

Question 1
It made me realise where Malgosia got her bright-eyed intelligence and her deep dark passion.

Is
a) her bright-eyed intelligence

| | a very good | a good | an acceptable | not a good |

b) her deep dark passion

| | a very good | a good | an acceptable | not a good |

description of Malgosia's character?

Notice that the choice of descriptive words to indicate the degree of prototypicality varied with different items.   For most items as in the example given

above, "good" was an appropriate description, but we also used "bad", "difficult" or "typical" to show more representative members, depending on the context. As in Task 3, the results obtained in Task 4 were correlated with the NS ratings obtained earlier.

4.4.3.5 Contributions of Tasks

Apart from being used as regular classroom practise to raise learners' awareness of the lexical knowledge they had obtained from the text, the four tasks also contributed to language learning and research investigation in various other ways.

In Task 1, Part 1 asked the L2 learners to write two examples that came to their minds under a given superordinate. This example-listing exercise related the current task to their personal experience and led to more effective involvement with their language learning. In Part 2, the four answer keys provided an opportunity for L2 learners to reflect on their language learning process, i.e. the reading process, and indicate whether they could recognize the items as having been read in the text. Furthermore, the researcher was interested to know whether the experience of reading might influence the ability to recognize the structure of semantic categories to any extent.

The open-ended nature of Task 2 provided information about the basis on which L2 learners understood the structure of the categories occurring in the text. In other words, we could glean an excellent idea of the extent to which the text was properly understood from the bracketted patterns made by the informants in order to accomplish the task. For instance, with an erroneous understanding of the text message, the informants might bracket semantically related items, rather than hierarchically structured ones, as the members of a category.

As for Task 3, though similar in demand to Task 2, it required learners to judge whether the underlined items could be regarded as members of the same category in the quoted context and hence consolidated the knowledge of what could be taken as a category member already obtained from reading the text.

In the final phase, a judgement on the degree of membership for each category member was required by Task 4. The learners therefore needed to recall the plot and setting of the reader that they had read. The overall process involved a reconstruction of the semantic categories present in the text, the recognition of potential category members and finally an assignment of prototypicality to each member within a category. The close association of reading comprehension and vocabulary depth learning, when defined as the ability to differentiate degrees of membership, is therefore acknowledged in these tasks.

4.5 Conclusion

This chapter began with a restatement of the particular perspective on vocabulary knowledge adopted in this research, including the importance of the contextual nature of semantic classification in second language acquisition contexts involving more advanced learners. Five hypotheses were constructed to test the depth of vocabulary knowledge of informants in terms of their ability to recognize semantic categories in a specific discourse context. In particular, two aspects of vocabulary knowledge, inclusion and membership, were under investigation. Rejecting a laboratory-like experimental design, the researcher decided to conduct the experiment in a normal classroom setting with the least possible disruption of regular practice. This involved selection of suitable pedagogic material for the would-be informants and negotiation with the regular teacher. A graded reader published by CUP was therefore selected, both because this type of pedagogic

material was currently used in the classroom and because the discourse context contained rich category structures ideal for the special characteristics of this research.

This research set out to investigate the properties of inclusion and membership of semantic categories in an extended text. The first step in identifying items suitable for testing the informants' ability to recognize category structure required the researcher to identify sets of category members. Inevitably, these sets of category members were sometimes accompanied by superordinate category names and sometimes not. This procedure was repeated and checked until the researcher was confident that unambiguous instances only had been collected. Great efforts were also made to ensure that each item was suitable for the hypothesis that was going to be investigated.

During the process of experiment design, difficulties arose with respect to identifying tests that would serve both as instruments of measurement and as practical regular classroom tasks. In addition, we did not find any suitable tests for vocabulary depth knowledge as defined in this research in the literature. As a consequence, self-invented tasks were the only option for the researcher. Drawing on the format of the existing vocabulary measurement instruments, the researcher designed four types of tasks to measure different aspects of the learner's ability to recognize the structure of semantic categories present in the context of the graded reader. Moreover, to fulfil the purpose of being used as tests, these tasks contributed to language learning and research investigation in various respects, and therefore provided valuable data capable of revealing the relationship between vocabulary learning, especially the acquisition of depth knowledge seen as context constrained semantic categorization, and the form in which this vocabulary was presented to the informants.

Before the implementation of the research, negotiation was required as to the practical constraints inherent in a real classroom setting. Of primary concern was the amount of time allowed for peripheral teaching materials such as readers in the context in which the research was conducted, and specifically how much reading at a time could be assigned to the informants and how the tasks could be conducted within the time constraints of the classroom. As a result of this limitation, the researcher had to rely on the language teacher to check comprehension of the reader and to train the informants to do the tests as required. Nevertheless, thanks to the cooperation of the teacher, data of a sufficient quality to be subjected to statistic analysis were collected. One important goal of this research, namely to disrupt the classroom practice as little as possible, was therefore accomplished.

Based on the data obtained, statistical analysis will be presented in the two following chapters, which reveals important information as to the informants' ability to recognize category structure. The results will be interpreted in light of vocabulary learning defined in terms of the ability to recognize the inclusion and membership properties of the semantic categories present in an extended text.

Chapter 5   Data Analysis and Results: Inclusion

In this chapter, the results of the experiments will be discussed in relation to the ability to recognize the inclusion properties of semantic categories.   It is assumed that an author constructs a text in such a way that the inclusion properties of lexical items will be maximally transparent to readers and readily absorbed and retrieved by them (see section 4.2.1).   Two kinds of text are used to contrast with the author's original text, text modified so that inclusion is encoded in a more structured way (see p. 111) and text modified so that inclusion is encoded in a more paratactic way (see p. 113).   Vocabulary learning is assessed in terms of the ability to recognize inclusion by means of Task 1.

The first section of the chapter will report the results for Hypothesis 1, which investigates the extent to which there is an association between a more structural encoding of an extended discourse text and successful vocabulary learning defined in terms of the ability to recognize the inclusion properties of semantic categories present in the text.   Data obtained from test items drawn from chapter 3 of the reader are presented in Task 1.   These data will be analyzed statistically and the results will be presented in tables.   After the analysis, conclusions and implications will be drawn in relation to pedagogical issues.

The second section of the chapter will focus on the investigation of Hypothesis 2, which examines a text modified so as to enable us to investigate the effect on vocabulary learning of a more paratactic encoding.   The data are provided by test items extracted from chapter 4 of the reader, also presented in Task 1, and will be submitted to the same statistical analysis as those conducted on the data provided by test items drawn from chapter 3.   By combining the results obtained from these

two sets of experimental data, we can draw conclusions with practical implications for the selection of pedagogic materials.

5.1 Experiment 2a, Inclusion and Structural Encoding

In this section, we will discuss the results of Task 1 in relation to Hypothesis 1. After a brief reiteration of Hypothesis 1 and Task 1, statistical analyses will be presented in relation to vocabulary learning defined in terms of the ability to recognize the inclusion of semantic categories based on the Task 1 scores. The data will be analyzed from different perspectives in order to provide a comprehensive set of results. Finally a discussion will be provided and several implications will be identified and discussed.

5.1.1 Hypothesis 1 and Task 1

Hypothesis 1

A text modified so that the inclusion properties of semantic categories are structurally encoded to a greater degree than by the original author will lead to less successful vocabulary learning (= less understanding of inclusion properties) than the author's original text.

Hypothesis 1 was designed to test the association of discourse structure and successful vocabulary learning in Experiment 2a. Specifically, a text in which superordinate category names and category members were encoded paratactically was modified so that the category names and members were encoded structurally. This was in order to enable the researcher to investigate whether the original text or the modified version would lead to more successful vocabulary learning. It was

assumed that the author's original text which conveyed depth meanings by means of paratactic arrangement would lead to more successful vocabulary learning than a modified text in which the same depth meaning was structurally encoded.

Hypothesis 1 was tested by Task 1, which contained two parts. Part 1 (Experiment 1) gave the informants a number of superordinate terms and required them to state whether they understood their meanings or not. In order that the researcher was able to confirm the reliability of the self-report, the informants were also asked to provide two examples of category members. Part 2 (Experiments 2a and 2b) required the informants to indicate whether or not the examples had been read and whether or not they could be included within the category whose superordinate name was given. Four answer keys were provided (A, B, C and D) and the informants were instructed to choose one. These keys indicated whether A) the informant thought s/he had read the item and it could be included in the superordinate category name; B) the informant thought s/he had read the item and it could not be included in the superordinate category name; C) the informant thought s/he had not read the item and it could be included in the superordinate category name; D) the informant thought s/he had not read the item and it could not be included in the superordinate category name.

5.1.2 Data Analysis

The informants were divided into two subgroups, the control and the experimental groups. The former read the original text and the latter read the modified version. Both groups took the same test (Task 1) and their answers were scored in the same way. Since key A indicated that the informants could recognize the exemplars as having been read and as members of the category indicated by the

superordinate category name provided, i.e. the inclusion properties were successfully

recognized by means of reading, key A answer was scored as 1 and all other answers

were scored as 0. The scores obtained from the informants in the control and the

experimental groups were then calculated and compared. The descriptive statistical

analyses are shown in Table 5.1, including means, standard deviations (SD) and the

mean difference between the two groups.

Table 5.1: Recognition of inclusion as a function of discourse context encoding: means and standard deviations of scores for all informants.

	n	A		
		Mean	SD	Mean difference
Control	14	10.93	4.92	-.435
Experimental	11	11.36	2.80	

Since Table 5.1 is derived from the scores of all the informants who were able

to recognize that they had read the item and considered that it was a member of the

superordinate category, its reliability may be compromised by virtue of including the

A scores for answers where informants were shown not to understand the

superordinate category names in Experiment 1. Therefore, the descriptive statistics

were replicated taking into account only those answers obtained from Experiment 2a

where informants had been shown to be aware of the meanings of the superordinate

category names in Experiment 1. In other words, we disregarded the answers

obtained from Experiment 2a where the informants stated in Experiment 1 that they

did not know the meanings of the superordinate category names when presented as

decontextualized items. The results are presented in Table 5.2.

Table 5.2:   Recognition of inclusion as a function of discourse context encoding: means and standard deviations of scores for items provided by informants aware of the meanings of the superordinate category names.

	n	A		
		Mean	SD	Mean difference
Control	14	9.50	4.64	.409
Experimental	11	9.09	3.51	

From these two Tables, we can see that the mean difference is inconsistent in that, in Table 5.1, the mean difference is a negative value, indicating that the experimental group outperforms the control group, whereas in Table 5.2 the mean difference is positive, indicating that the control group outperforms the experimental group.

In order to determine the possible significance of these results, the means between the control and the experimental groups were then compared, using an independent samples t-test.   As shown in the last column in Table 5.3, the difference does not reach a level of significance ($p > .05$).

Table 5.3:   Independent Samples t-Test for recognition of inclusion as a function of discourse context encoding for all informants.

	df	t	Sig.
Control (14) Experimental (11)	23	-.261	.398

That is, we have to accept the null hypothesis that the modified text does not

lead to more successful vocabulary learning[1]. Therefore, we can only say that

adapting a text so that inclusion is encoded more structurally than by the original

author does not affect recognition of inclusion. Put another way, it seems that the

more structured text that results from bringing semantically related items closer

syntactically does not make awareness of their semantic interrelation more salient.

The effect size was also computed in order to show whether there was any

possible trace of an effect. The results are shown in Table 5.4.

Table 5.4: Effect size for recognition of inclusion as a function of discourse context encoding for all informants.

	ES (Effect Size)	ES on control group SD
Control (14) Experimental (11)	0.10	0.09

Two values are given in Table 5.4, one based on the pooled standard deviation

and the other on the standard deviation of the control group. This is because there

is an ongoing debate regarding which SD (standard deviation) should be used in

calculating effect size (Coe, 2002). As the Table shows, an effect size of 0.10 was

obtained using a pooled standard deviation and 0.09 using the standard deviation of

the control group. Both show only a very small effect. A repeated procedure was

conducted on the means and SD of the scores for the informants aware of the

meanings of the superordinate category names. The results are presented in Table

5.5 and the effect size in Table 5.6.

---

[1] "Null hypothesis" is not used here in the strict sense in which it is commonly used, i.e. to refer to results which are the reverse of those predicted by an experimental hypothesis. Here, we simply meant that no difference was found for the learning outcomes of the control and experimental groups.

163

Table 5.5: Independent Samples t-Test for recognition of inclusion as a function of discourse context encoding for items provided by informants aware of the meanings of the superordinate category names.

	df	t	Sig.
Control (14) Experimental (11)	23	.243	.405

Table 5.6: Effect size for recognition of inclusion as a function of discourse context encoding for items provided by informants aware of the meanings of the superordinate category names.

	ES (Effect Size)	ES on control group SD
Control (14) Experimental (11)	-0.10	-0.09

Table 5.5 shows the independent samples t-test results for items where the informants were aware of the meanings of the superordinate category names. Again, the result does not reach a level of significance. In addition, the effect size in Table 5.6 shows a negative value, contrary to the result for all the informants as presented in Table 5.4. We therefore conclude that the results not only fail to attain any level of significance but that the results are inconsistent.

5.1.3 Additional Analysis

In addition, the frequency of each key chosen by all informants for all questions was computed with a view to seeing whether there would be a difference between the two groups with respect to the effect of modification on recognition of having read an item and on the recognition of inclusion regardless of reading effect. Table 5.7 shows the frequency of each key chosen by all informants in the control

and the experimental groups for all questions, while Table 5.8 shows the frequency of each key chosen by the informants in the control and the experimental groups who were aware of the meanings of the superordinate category names. In other words, the frequencies in Table 5.8 exclude those of the informants who stated that they did not understand the superordinates when presented out of context in Experiment 1.

Table 5.7: Frequency of each key chosen by all informants.

Group/Key	A	B	C	D
Control	153 (45.8%)	55 (16.4%)	64 (19.2%)	62 (18.6%)
Experimental	125 (47.3%)	41 (15.5%)	44 (16.7%)	54 (20.5%)

Table 5.8: Frequency of each key chosen by informants aware of the meanings of the superordinate category names.

Group/Key	A	B	C	D
Control	133 (47.3%)	49 (17.4%)	48 (17.1%)	51 (18.2%)
Experimental	100 (48.3%)	29 (14.0%)	35 (16.9%)	43 (20.8%)

Since both key A and B responses indicate that the informants recognize the examples as having been read, it was decided to determine whether the modification had any impact on recognition of having read an item. Therefore, the frequency of A and B responses were aggregated and the means and standard deviations of both the control and the experimental groups analyzed, as shown in Table 5.9.

Table 5.9:   Recognition of having read an item as a function of discourse context encoding (=reading effect): means and standard deviation for all informants.

	n	AB		
		Mean	SD	Mean difference
Control	14	14.86	5.40	-.234
Experimental	11	15.09	3.91	

The mean difference is negative, indicating that the experimental group outperformed the control group in recognizing test items as having been read. Nevertheless, this result is not replicated when we only take into account the response of the informants aware of the meanings of the superordinate category names, as shown in Table 5.10.

Table 5.10:   Recognition of having read an item as a function of discourse encoding (=reading effect): means and standard deviation for informants aware of the meanings of superordinate category names.

	n	AB		
		Mean	SD	Mean difference
Control	14	13.00	5.45	1.272
Experimental	11	11.73	4.05	

The mean difference in this table has a positive value, indicating that the control group outperformed the experimental group in recognizing test items as having been read.   Accordingly, no definite conclusion can be drawn based on these conflicting results.

An independent samples t-test was also conducted and the effect size calculated. Table 5.11 shows that the difference between the control and the experimental

groups is not statistically significant (*p*> .05) and the effect size in Table 5.12 is very

small, irrespective of which standard deviation is used.

Table 5.11:   Independent Samples t-Test for reading effect for all informants.

	df	t	Sig.
Control (14) Experimental (11)	23	-.121	.452

Table 5.12:   Effect size of reading effect for all informants.

	ES (Effect Size)	ES on control group SD
Control (14) Experimental (11)	0.05	0.04

Table 5.13 displays the results for informants aware of the meanings of the

superordinate category names, and Table 5.14 shows the results of effect size.

*Table 5.13:   Independent Samples t-Test for reading effect for informants aware of
the meanings of the superordinate category names.*

	df	t	Sig.
Control (14) Experimental (11)	23	.646	.262

Table 5.14: Effect size for reading effect for informants aware of the meanings of the superordinate category names.

	ES (Effect Size)	ES on control group SD
Control (14) Experimental (11)	-0.26	-0.23

In Table 5.14, the effect size is negative (ES= -0.26 & ES= -0.23), preventing us from drawing any conclusion associating modification and the awareness of having read an item.   In other words, at this level at least, the degree of structural encoding has no effect on awareness of reading.

The same procedure was applied to key A and C responses since both indicated recognition of inclusion.   Means and standard deviations for these results are shown in Tables 5.15 and 5.16, with Table 5.15 representing the aggregated results obtained from all informants whether or not they were aware of the meanings of the superordinate category names, and Table 5.16 the aggregated results obtained only from informants aware of the meanings of the superordinate category names.

Table 5.15: Recognition of inclusion regardless of reading effect: means and standard deviations for all informants.

	n	AC		
		Mean	SD	Mean difference
Control	14	15.50	3.52	-.409
Experimental	11	15.91	3.36	

Table 5.16:   Recognition of inclusion regardless of reading effect: means and standard deviations for informants aware of the meanings of the superordinate category names.

	n	AC		
		Mean	SD	Mean difference
Control	14	12.93	4.03	.656
Experimental	11	12.27	3.64	

As can be seen, the mean difference in Table 5.15 is a negative value, indicating that the experimental group outperformed the control group, while in Table 5.16 the value is positive, indicating that the control group outperformed the experimental group.   Therefore, no conclusion can be drawn from these contradictory results.

Finally, an independent samples t-test and an effect size calculation were also performed.   No significant difference was found between the two groups with respect to recognition of inclusion for all informants ($p>$ .05) (Table 5.17) or for informants aware of the meanings of the superordinate category names only (Table 5.19).   The effect sizes are presented in Tables 5.18 and 5.20

Table 5.17:   Independent Samples t-Test for recognition of inclusion regardless of reading effect for all informants.

	df	t	Sig.
Control (14) Experimental (11)	23	-.294	.385

Table 5.18: Effect size for recognition of inclusion regardless of reading effect for all informants.

	ES (Effect Size)	ES on control group SD
Control (14) Experimental (11)	0.12	0.12

Table 5.19: Independent Samples t-Test for recognition of inclusion regardless of reading effect for informants aware of the meanings of the superordinate category names.

	df	t	Sig.
Control (14) Experimental (11)	23	.421	.338

Table 5.20: Effect size of recognition of inclusion regardless of reading effect for informants aware of the meanings of the superordinate category names.

	ES (Effect Size)	ES on control group SD
Control (14) Experimental (11)	-0.17	-0.16

The effect size for all informants in Table 5.18 is 0.12 standard deviation units and –0.17 in Table 5.20 for those informants aware of the meanings of the superordinate category names. Again, the results are inconsistent and no conclusion can be drawn.

5.1.4 Discussion

None of the analyses mentioned in sections 5.1.2 and 5.1.3 reveals any

significant difference between the control and the experimental groups. Indeed, if anything, they present an ambiguous picture in which the results obtained from all the informants contradict those obtained only from informants aware of the meanings of the superordinate category names. In Hypothesis 1, it was assumed that vocabulary learning would be less successful when the author's original text was modified in such a way as to bring semantically related superordinate category names and category members into a single structural unit. No evidence can be found in support of this hypothesis. Even when effect size is computed, the value is very small (ES= 0.10), indicating that the effect of modification is barely noticeable. A contradictory result occurs when we take into account only the scores of informants aware of the meanings of the superordinate category names. In this case, the control group slightly outperforms the experimental group and the effect size becomes negative. We therefore have to accept the null hypothesis that a more structural encoding than that chosen by the original author does not make any difference to successful vocabulary learning defined in terms of the ability to recognize the inclusion properties of semantic categories. A potential explanation was mentioned in the literature review regarding vocabulary learning that, since a greater degree of structural encoding made the message more readily processable, less effort was required and consequently, no deeper traces were created to enable the informants to retain these items in long-term memory.

The "no effect" findings in relation to Hypothesis 1 are repeated when the data are considered from a range of other perspectives. If we assume that the effect of reading is not so pronounced and compute the frequency of A and C responses for all informants, i.e. if we compute successful recognition of inclusion regardless of reading effect, it turns out that the control group outperforms the experimental group.

However, this result is reversed in the case of informants aware of the meanings of the superordinate category names. Under these circumstances, we are left with the inconclusive finding that enhanced structural encoding is neither more nor less likely to lead to successful vocabulary learning.

For the time being, therefore, no conclusive result can be found with respect to the relationship between the structuration of a discourse context and the success of the kind of deep vocabulary learning tested for. One conclusion that might be drawn is that the degree of structuration of a reading text is irrelevant and has no significance for vocabulary learning of this kind. At the same time, it should be noted that the modified text connects the concepts expressed by the superordinate category names and their members in a single sentential unit, and the length of the sentence is thus increased. An earlier study showed that a passage which maintained surface clues to understanding a semantic relationship, such as the structures of relative pronoun and subordinate clause, might increase comprehensibility because the semantic relationship between components was unambiguously stated (Blau, 1982). However, as our results show, this finding is not replicated for inclusion.

Another possible factor to consider is the level of proficiency of the informants. The proficiency level of language learners is always an important variable in empirical research. In research cited in chapter 2, the same treatment produced different effects on learners of different proficiency levels (Hulstijn, 1993; Knight, 1994; Prince, 1996). Regrettably, due to the particular characteristics of this research, this variable was kept constant. One recommendation for any future study would be to test for any effect of this variable.

From another point of view, since the present research relates to recognition of

inclusion only, structuration may be important in other aspects of vocabulary learning, such as breadth learning. A more structured text may help or impede the learner's ability to increase vocabulary size. Because a more structured text relies on structural dependencies and thus consists of longer sentences, it may require more resources to decode. Therefore, a lexical item may not be so easily assimilated, thus cancelling out any positive effect of close structuration and the contextual clues longer sentences provide.

Several implications may be drawn from the present findings. First, for the selection of pedagogic materials, especially for reading courses, criteria other than structuration should be taken into consideration to improve vocabulary learning. It is not for us to speculate here on just what these criteria might be, but they might include the degree of syntactic complexity, with longer and more complex sentences expected to hinder vocabulary learning; the degree of difficulty of the lexical items encountered, such as infrequent words used in a text possibly requiring vocabulary instruction; and the grouping of lexical items in semantic or thematic clusters and the possible effect on retention. Moreover, a sensitivity to learner differences is also required in selecting an appropriate reading text, since materials at the same pedagogic level may produce different learning outcomes for different learners.

In order to intervene in normal classroom practice as little as possible, the researcher did not test the level of reading comprehension and hence was unable to investigate any possible relationship between level of reading comprehension and acquisition of vocabulary depth knowledge when defined as the ability to recognize inclusion. In the literature, it was found that vocabulary knowledge (both breadth and depth) was associated with level of reading comprehension (see section 2.4.4). For example, Laufer investigated the relationship between vocabulary size and

reading comprehension by using Nation's Vocabulary Levels Test and Meara & Jones' Eurocentres Vocabulary Test to measure vocabulary size and reading comprehension by two different reading tests (Laufer, 1992). High correlations were found between vocabulary size and reading comprehension. Additional research into the interrelationship between vocabulary size and depth and their relationship with reading comprehension also revealed a high correlation as measured by the Vocabulary Levels Test and a reading comprehension test adopted from TOEFL (Qian, 1999). It also shows a relationship between depth of vocabulary knowledge as measured by the Word Associate Format devised by Read (1993) and reading comprehension (Qian, 1999).

Nevertheless, these studies are not directly comparable to the present research because vocabulary knowledge in this research is specifically defined in terms of the ability to recognize the inclusion properties of semantic categories, whereas previous studies define vocabulary breadth traditionally in terms of recognition of single words and vocabulary depth in terms of word associates.

For materials developers, and especially for authors of classic and original readers, an over-structured text may not produce the desired learning outcome. Based on the results obtained, what the researcher shows is that when the writer's default is paratactic, the structured option is neither better nor worse. This is somewhat surprising. It seems that for an extended discourse text such as a graded reader, there is no best way to present a text with respect to the variable tested for here. Thus, successful reading will depend on other variables, such as proficiency level.

In this section of the chapter, we have investigated only one aspect of discourse structure and found no conclusive results with respect to the hypothesis tested. In

the next section, the results obtained with respect to vocabulary learning when a more structurally encoded text was modified so as to render the encoding more paratactic will be discussed and compared to the results discussed in this section.

5.2 Experiment 2b, Inclusion and Paratactic Encoding

We also perceive a possibility that modification away from structural encoding toward more paratactic encoding may have an effect on more successful vocabulary learning. Therefore, in this section, we will discuss the effect on vocabulary learning defined in terms of the ability to recognize the inclusion properties of semantic categories as stated in Hypothesis 2. Data obtained in Experiment 2b will be submitted to statistical analyses identical to those performed in the previous section. In particular, the results and findings will be compared to what we have found in relation to Hypothesis 1. One way or another, this comparison may be expected to provide a more conclusive result.

5.2.1 Hypothesis 2 and Task 1

Hypothesis 2

A text modified so that the inclusion properties of semantic categories are paratactically encoded to a greater degree than by the original author will lead to less successful vocabulary learning (= less understanding of inclusion properties) than the author's original text.

Hypothesis 2 was designed to test the association of another aspect of discourse structure and successful vocabulary learning. In contrast to Experiment 2a, a text in which superordinate category names and category members were originally encoded

structurally was modified so that the category names and members were encoded more paratactically in order to enable the researcher to investigate whether the original author's more structural encoding or the modified version would lead to more successful vocabulary learning. In Hypothesis 2, it was assumed that the author's original text, i.e. the more structural encoding which conveyed depth meanings by means of structural relation, would lead to more successful vocabulary learning than a modified text in which the same depth meaning was paratactically encoded.

Just as with Hypothesis 1, Hypothesis 2 was tested by Task 1, which contained test items taken from chapter 3 and chapter 4 of the reader. Items from chapter 3 were used to test Hypothesis 1 (Experiment 2a), whereas those from chapter 4 were used to test Hypothesis 2 (Experiment 2b). The informants' answers obtained from Experiment 2b were taken as essential data in order to test Hypothesis 2, their reliability was established in relation to recognition data provided in Experiment 1.

5.2.2 Data Analysis

As in Experiment 2a, the class was divided into two subgroups, the control and the experimental groups. However, the group which read the modified text of chapter 3 in Experiment 2a read the original version of chapter 4, and hence was the control group; and the group which read the original text of chapter 3 in Experiment 2a read the modified version of chapter 4, and hence was the experimental group. Both groups took the same test, Task 1. The scoring process was the same as that described in section 5.1.2, i.e. a key A answer was given a score of 1 and all the other answers were scored 0 because key A showed that the informants could recognize the item as included in the category, an effect resulting from reading

experience.   The scores obtained from the control and the experimental groups were then calculated and compared.   The descriptive statistical analysis is shown in Table 5.21, including means, standard deviations and mean difference between the two groups.

Table 5.21:   Recognition of inclusion as a function of discourse context encoding: means and standard deviations of scores for all informants.

	n	A		
		Mean	SD	Mean difference
Control	11	6.82	2.36	-.682
Experimental	14	7.50	3.03	

Since Table 5.21 is derived from the scores of all the informants, its reliability may also be compromised by including the scores for answers where informants were shown not to understand the superordinate category names presented as isolated items in Experiment 1.   Therefore, the statistical analysis was repeated taking into account only those answers in Experiment 2b where informants had been shown to be aware of the meanings of the superordinate category names in Experiment 1.   The results are presented in Table 5.22.

Table 5.22:   Recognition of inclusion as a function of discourse context encoding: means and standard deviations of scores for items provided by informants aware of the meanings of the superordinate category names.

	n	A		
		Mean	SD	Mean difference
Control	11	5.82	2.44	-.682
Experimental	14	6.50	2.62	

These two tables clearly indicate that the mean difference is consistent, both being a negative value; this is very different from the previous findings.   Both tables indicate that the experimental group outperformed the control group and hence that paratactic modification might have an effect.

To determine whether or not the effect reached significance, an independent samples t-test was performed to compare the means between these two groups.   The results are presented in Table 5.23.

Table 5.23:   Independent Samples t-Test for recognition of inclusion as a function of discourse context encoding for all informants.

	df	t	Sig.
Control (11) Experimental (14)	23	-.613	.273

As shown in Table 5.23, however, the difference does not reach a level of significance ($p > .05$).   Therefore, we have to accept the result that a more paratactic encoding of the text does not lead to more successful vocabulary learning.   That is to say, the modified text in which the superordinate category names and the category members were presented paratactically does not make a significant difference to the

recognition of inclusion.

The results of calculating effect sizes are shown in Table 5.24, so as to show any possible trace of an effect.

Table 5.24: Effect size for recognition of inclusion as a function of discourse context encoding for all informants.

	ES (Effect Size)	ES on control group SD
Control (11) Experimental (14)	0.25	0.29

The figure in the second column is based on the pooled standard deviation and the final column on that of the control group. Both are shown to exceed the value of 0.2, which is considered a small effect but noticeable effect. In his definition, Cohen classified 0.20 standard deviations as a small effect, 0.50 a medium effect, and 0.80 a large effect (Cohen, 1969: 23-4). Though the effect found was small, it is worth noting that modification by means of paratactically encoding the superordinate category names and the category members does produce a small trace effect. The same calculating procedures were repeated on the means and standard deviations of scores for the informants aware of the meanings of the superordinate category names. The results are shown in Table 5.25 and the effect size shown in Table 5.26.

Table 5.25:   Independent Samples t-Test for recognition of inclusion as a function of discourse context encoding for items provided by informants aware of the meanings of the superordinate category names.

	df	t	Sig.
Control (11) Experimental (14)	23	-.665	.256

Table 5.26:   Effect size for recognition of inclusion as a function of discourse context encoding for items provided by informants aware of the meanings of the superordinate category names.

	ES (Effect Size)	ES on control group SD
Control (11) Experimental (14)	0.27	0.28

Although the result of an independent samples t-test does not show any significant difference, the effect sizes are consistent with those in Table 5.24, indicating that a small but traceable effect occurs.   We are therefore able to conclude that the modified text produces an effect on vocabulary learning, though the effect is small and does not reach a level of significance.   In other words, the modified text in which the superordinate category names and the category members are paratactically encoded might have a small effect on the recognition of inclusion of the semantic categories present in the text.   These results are surprising but very important when compared with the findings in the previous section in that they imply that a more paratactic encoding of discourse context might affect vocabulary learning to some extent.

5.2.3 Additional Analysis

As in Experiment 2a, the frequency of each key chosen by all informants for all questions was computed with a view to seeing whether there would be a difference between the two groups with respect to the effect of modification on recognition of having read an item and on the recognition of inclusion regardless of any reading effect. Table 5.27 shows the frequency of each key chosen by all informants in the control and the experimental groups for all questions, while Table 5.28 shows the frequency of each key chosen by the informants in the control and the experimental groups who were aware of the meanings of the superordinate category names.

Table 5.27: Frequency of each key chosen by all informants.

Group/Key	A	B	C	D
Control	75 (52.4%)	20 (14.0%)	21 (14.7%)	27 (18.9%)
Experimental	89 (62.7%)	18 (12.6%)	21 (14.8%)	14 (9.9%)

Table 5.28: Frequency of each key chosen by informants aware of the meanings of the superordinate category names.

Group/Key	A	B	C	D
Control	64 (53.4%)	16 (13.3%)	18 (15.0%)	22 (18.3%)
Experimental	77 (65.2%)	12 (10.2%)	16 (13.6%)	13 (11.0%)

In order to determine whether the modification had any impact on recognition of having read an item, the frequency of A and B responses was aggregated because both responses indicate that the informants could recognize the examples as having

been read. Means and standard deviations of both groups were analyzed and are

shown in Table 5.29.

Table 5.29: Recognition of having read an item as a function of discourse context encoding (=reading effect): means and standard deviations for all informants.

	n	AB		
		Mean	SD	Mean difference
Control	11	8.64	2.84	-.435
Experimental	14	9.07	3.34	

The mean difference in Table 5.29 shows a negative value, indicating that the

experimental group outperformed the control group in recognizing test items as

having been read.   To see whether the same result holds for informants aware of the

meanings of the superordinate category names, the calculating procedures were

repeated and the results are shown in Table 5.30.

Table 5.30:   Recognition of having read an item as a function of discourse encoding (=reading effect): means and standard deviations for informants aware of the meanings of superordinate category names.

	n	AB		
		Mean	SD	Mean difference
Control	11	7.27	2.94	-.370
Experimental	14	7.64	2.59	

As shown in Table 5.30, the mean difference is also a negative value,

consistent with the result in Table 5.29.   This actually indicates that the

experimental group outperformed the control group in recognizing test items as

having been read and possibly that the modified text might have an effect on the recognition of having read an item.

To see whether the effect of modification reached a level of significance, an independent samples t-test was conducted and the results are presented in Table 5.31. Additionally, the effect size was calculated to look for any trace effect. The results are presented in Table 5.32.

Table 5.31:   Independent Samples t-Test for reading effect for all informants.

	df	t	Sig.
Control (11) Experimental (14)	23	-.345	.366

Table 5.32:   Effect size for reading effect for all informants.

	ES (Effect Size)	ES on control group SD
Control (11) Experimental (14)	0.14	0.15

As shown in Table 5.31, the difference between the control and the experimental groups is not statistically significant. The effect size in Table 5.32 is smaller than 0.20 irrespective of which standard deviation is used. Both an independent samples t-test and effect size calculation were repeated with the data for informants aware of the meanings of the superordinate category names and results are presented in Tables 5.33 and 5.34.

Table 5.33: Independent Samples t-Test for reading effect for informants aware of the meanings of the superordinate category names.

	df	t	Sig.
Control (11) Experimental (14)	23	-.335	.370

Table 5.34: Effect size for reading effect for informants aware of the meanings of the superordinate category names.

	ES (Effect Size)	ES on control group SD
Control (11) Experimental (14)	0.13	0.13

In addition to the non-significant t-test result, the effect size is consistent with that in Table 5.32, i.e. a smaller effect than 0.20. At this point, we can conclude that the effect of a modified text on recognition of having read an item is not statistically significant. In fact, the effect size is very small indeed. Therefore, paratactic modification does not increase awareness of having read items treated as category members.

The same procedure was applied to key A and C responses to determine any effect that the modified text might have on recognition of inclusion regardless of reading effect. The means, standard deviations and mean differences are shown in Tables 5.35 and 5.36. Table 5.35 gives the results obtained from all informants whether or not they indicated awareness of the meanings of the superordinate category names, whereas Table 5.36 includes those obtained only from informants aware of the meanings of the superordinate category names.

Table 5.35: Recognition of inclusion regardless of reading effect: means and standard deviations for all informants.

	n	AC		
		Mean	SD	Mean difference
Control	11	8.73	1.90	-1.201
Experimental	14	9.93	1.59	

Table 5.36: Recognition of inclusion regardless of reading effect: means and standard deviations for informants aware of the meanings of the superordinate category names.

	n	AC		
		Mean	SD	Mean difference
Control	11	7.45	2.50	-.760
Experimental	14	8.21	2.00	

Both tables show a negative value of mean difference, indicating that the experimental group outperformed the control group. The larger mean difference shown in Table 5.35 based on the scores from all the informants might be taken to suggest that the modified text has an effect on recognition of the inclusion regardless of recognition of having read an item. Therefore, an independent samples t-test and effect size calculation were conducted. The results of independent samples t-test are shown in Table 5.37, and the effect size shown in Table 5.38.

Table 5.37:   Independent Samples t-Test for recognition of inclusion regardless of reading effect for all informants.

	df	t	Sig.
Control (11) Experimental (14)	23	-1.720	.0495

Table 5.38:   Effect size for recognition of inclusion regardless of reading effect for all informants.

	ES (Effect Size)	ES on control group SD
Control (11) Experimental (14)	0.69	0.63

As shown in Table 5.37, a marginally significant difference is found between the two groups with respect to recognition of inclusion for all informants (p= .05) and in Table 5.38 there is a distinct effect size of 0.63, a value larger than the medium value of 0.50 but smaller than the large value of 0.80 (Cohen, 1969). This enables us to conclude that the effect of the modified text on the recognition of inclusion regardless of reading effect can reach significance, though it is marginal. Put another way, the modification of a text toward more paratactic encoding might lead to more successful vocabulary learning defined in terms of the ability to recognize inclusion.

To ensure that this result extends to the scores for the informants aware of the meanings of the superordinate category names, an independent samples t-test was repeated on these scores and the results are shown in Table 5.39. The results of testing for effect size are shown in Table 5.40.

Table 5.39: Independent Samples t-Test for recognition of inclusion regardless of reading effect for informants aware of the meanings of the superordinate category names.

	df	t	Sig.
Control (11) Experimental (14)	23	-.843	.204

Table 5.40: Effect size for recognition of inclusion regardless of reading effect for informants aware of the meanings of the superordinate category names.

	ES (Effect Size)	ES on control group SD
Control (11) Experimental (14)	0.34	0.30

As shown here, the p value does not reach a level of significance and the effect size is not as great as that in Table 5.38 for all the informants. Still, it is larger than the small value of 0.20 standard deviations but smaller than the medium value of 0.50. Therefore, we can only say that there is an effect of modification on the recognition of inclusion regardless of reading effect, but it is small. In other words, a more paratactic encoding of superordinate category names and category members might have an effect on vocabulary learning defined in terms of the ability to recognize the inclusion of semantic categories.

5.2.4 Discussion

The analysis of the data obtained in Experiment 2b reveals a very different picture from the analysis of the data obtained in Experiment 2a. Although in many cases the statistical analyses did not reach significance, the results were more

consistent. A noticeable and positive effect could always be detected regarding the ability to recognize inclusion whether it was based on the scores of all informants or on the scores of the informants aware of the meanings of the superordinate category names when presented as discrete items. Furthermore, this effect was shown to reach a level of marginal significance in the analysis of A and C responses, disregarding whether or not the informants recognized the items as having been read.

In Hypothesis 2, it was predicted that successful vocabulary learning would be less likely to occur when the author's original text was modified in such a way as to represent semantically related superordinate category names and category members in a more paratactic structure. We found an effect in the effect size of A responses (A stating that the informant thought s/he had read the item and it could be included in the superordinate category name) whether or not the informants were aware of the meanings of the superordinate category names, and stronger evidence in the scores of A and C responses for all informants (C stating that the informant thought s/he had not read the item and that it could be included in the superordinate category name). The positive effect consistently spread to every analysis with respect to the recognition of inclusion. Such a result was very surprising, i.e. not predicted by the hypothesis, and even more so when compared to the findings in Experiment 2a for structural modification. In contrast to the previous findings, the evidence showed that when the author's default is structural encoding, the paratactic option increases awareness of superordinate category names and category members and enhances the ability to recognize inclusion. The finding surprises us because it disconfirms the prediction of Hypothesis 2 and reveals that more successful vocabulary learning occurs when the semantic categories of the text are paratactically encoded to a greater degree than in the author's original structurally encoded text, at least for this

group of learners and this text.

Hard as we tried, we found it difficult to explain why this was the case, except that a paratactic structure was considered to be more natural (Rynell, 1952: 46). But given that paratactic encoding in the author's original text did not produce a noticeable and definite effect on vocabulary learning (Experiment 2a), such an explanation might not be regarded as robust. Considering the present results with those obtained from Experiment 2a, we can only say that, surprisingly, the author does not achieve a maximally effective way of communicating his meanings. Other options can result in more successful transmission of meaning. The essential meaning particular to a specific reader (text not person) in this research is taken to consist in context-dependent inclusion, which defines vocabulary depth knowledge so that successful vocabulary learning entails the ability to recognize inclusion properties. Moreover, successful reading comprehension can be achieved only when the meaning of the text can be grasped, that is, the inclusion properties of semantic categories can be absorbed and retrieved. Thus, reading and vocabulary learning are closely related. That is to say, when lexical input is provided in an extended text, understanding/learning these lexical items requires the reader to understand the way they are related, i.e. the way they constitute semantic categories. Thus, the text provides the context which defines these as contextually determined semantic categories. Reading the text so as to understand the category relationship of items is the initial step to learning how these items are used. As reviewed in chapter 2, vocabulary learning from context has been found essential for vocabulary development. We find additional evidence in this research that reading contributes to vocabulary learning to a greater extent.

Linking vocabulary depth knowledge and reading in such a way will provide

insightful information for publishers interested in graded readers. As a regular practice, they instruct writers of graded readers as to the vocabulary size (breadth) permitted in a reader at a given level, but seem not to have reliable ways of measuring vocabulary depth knowledge. The present research suggests that vocabulary depth knowledge is important because it provides a more effective reading experience for advanced learners. In addition, the degree of structuration may be associated with understanding of vocabulary and hence the way writers create the categories that are particular to their texts.

For vocabulary learning, the particular point of view adopted by this research provides us with an opportunity to investigate the relationship between categorization and vocabulary learning, a relationship which is not yet widely acknowledged. It is believed that language learning is closely related to human cognition. As an essential aspect of cognitive function, categorization, is therefore assumed to be central to language and particularly to the mental lexicon and thus to vocabulary acquisition.

With reference to the selection of pedagogical materials, the results in relation to Hypothesis 2 imply that a text with more paratactic encoding than devised by the original author has the potential to enhance vocabulary learning. The different degrees of paratactic and structural encoding which have been shown here to influence vocabulary learning to some limited degree have so far been largely neglected in the literature. Although the actual mechanism by which this effect is achieved is still obscure, it is worth noting and deserves further investigation. Especially for anyone involved in materials development, such as writers of extended texts, the different options for encoding meaning need to be taken into account since they affect the communication of meaning and hence the effectiveness

of language learning.

5.3 Conclusion

In this chapter, we have investigated the association of successful vocabulary learning and different encoding options for an extended text. Given that successful vocabulary learning was defined in terms of the ability to recognize the inclusion property of semantic categories present in a text under study, we demonstrated that modification toward a more structured encoding did not lead to either more successful or less successful vocabulary learning, whereas modification toward more paratactic encoding produced a small but noticeable effect on vocabulary learning. This has important implications for the selection of pedagogical materials in that an over-structured text may not produce the desired learning outcomes while a more paratactic text may lead to more effective learning.

Moreover, this research has raised an important issue in relation to one aspect of vocabulary learning which has been largely ignored, i.e. the hierarchical structure of semantic categories invoked by the lexical items encountered by second language learners in a particularly defined context. The role of categorization in second language acquisition, and specifically in the acquisition of L2 vocabulary, was therefore investigated. In addition, by defining vocabulary depth knowledge in terms of the ability to recognize inclusion, the researcher was able to explore the importance of depth vocabulary learning in relation to reading comprehension.

In addition, cultural knowledge might play a role in recognizing inclusion properties of these contextually determined semantic categories. For instance, while judging whether *she had light brown hair* could be included in the category *Derek's reasons for liking Rachel*, the hair colour might be an influential criterion in

making the decision because normally Chinese people have black hair. The informants might like or dislike hair of a different colour and, consciously or unconsciously, take it as an important feature for personal likeness. As a consequence, certain hidden aspects of cultural knowledge probably affected the overall judgement of category membership, although we were not able to point out what they were and how they functioned in these tasks. Future studies of this kind may compare informants from different cultural backgrounds, such as Western and Oriental.

The limits of these experiments have also been discussed. That is, because a relatively advanced level of proficiency was required for the investigation of depth knowledge, the factor of learner proficiency level was held invariant in these experiments and was not therefore taken into account. The researcher suggests that for future research of this kind, proficiency level might be an intriguing variable deserving more attention, as has already been shown in a considerable number of studies on other aspects of vocabulary learning. Future research may therefore try to discriminate the effects of structuration in relation to degrees of advancedness.

Chapter 6   Data Analysis and Results: Membership

In this chapter, we discuss the results obtained in our investigation of category membership.   Specifically, two aspects of membership are investigated and the areas of investigation are formulated in three hypotheses, viz. Hypotheses 3, 4, and 5. In the first section of this chapter, we discuss the results obtained in relation to Hypothesis 3, which was designed to test whether or not the presence of category names enabled informants to recognize members of a category.   Successful vocabulary learning is thus defined in terms of the ability to recognize that a lexical item is a member of a particular category.   Data obtained from Task 2 are used to investigate this hypothesis and subjected to statistical analyses to determine whether there is any association between the presence of category names and the ability to recognize category members.

The second section of the chapter will discuss the extent of any association linking the ability to recognize membership and rhetorical patterning among category members, as stated in Hypothesis 4.   Two sets of data will be compared. One set of data is obtained from Task 3, in which second language learners are required to choose the correct answer from a multiple-choice question in which putative members of a common category are indicated by underlining.   The second set consists of ratings of degree of rhetorical coupling identified by native speakers on the same test items.   A statistical correlation test will be performed on these two sets of data to see whether there is an association between the ability to recognize category membership and the degree of rhetorical coupling exhibited by a set of members of a common category.

Finally, the third section presents data relating to Hypothesis 5, which is

designed to investigate whether there is an association between discrimination of degree of category membership and rhetorical coupling. Hypothesis 5 predicts that a lower degree of rhetorical coupling shared by category members will lead to a finer distinction by learners of the degree of category membership. As in section 2, two sets of data will be compared. One set consists of the data obtained in Task 4 measuring the performance of second language learners; the second consists of ratings of degree of rhetorical coupling identified by native speakers on the same test items. Distinct data sets are used to test Hypothesis 4 and Hypothesis 5 for the reasons given in 4.4.2.3 (p. 145). The two data sets obtained to test Hypothesis 5 will be subjected to the same correlational analysis as the data sets discussed in section 2 in order to test for any possible association between the ability of learners to discriminate degree of category membership and rhetorical coupling.

6.1 Experiment 3, Membership

In this section, we will analyze the data obtained in Task 2 to see whether there is an association between the presence of the superordinate category name or label and the ability to recognize category members. The analysis will be conducted separately on test items containing 2, 3, and 4 category members. In particular, we employ a paired-samples t-test to analyze each pair and to determine whether there is any significant difference within a pair. The results will also be discussed in relation to the research on categorization that has been conducted in experimental psychology.

6.1.1   Hypothesis 3 and Task 2

<u>Hypothesis 3</u>

Successful vocabulary learning (= recognition of category membership) is
more likely to occur where multiple instances of membership at the same level
in a semantic category occur alongside co-text which provides the category
label, i.e. indicates the level of inclusion of those instances.

Hypothesis 3 was designed to test whether the presence of superordinate
category names would enable informants to recognize the members of a category and
organize the complete internal category structure present in the text.   The ability to
recognize all the members of a category when the superordinate category name is
present in the co-text is contrasted with the ability to recognize an equal number of
category members as category members when the superordinate category name is
not present in the co-text.   It was assumed that the indication of the level of
inclusion as a result of the presence of the superordinate category name occurring in
the co-text would lead to successful vocabulary learning defined in terms of the
ability to recognize category membership.

Task 2 was designed to test Hypothesis 3.   It provided the entire discourse
context where a category occurred and required informants to identify all the
members within the category by bracketing them.   In other words, correctly putting
the boundary before and after all the category members was considered to indicate
successful recognition of members.   The number of category members was
explicitly stated for each question and the questions were arranged so that first
two-member categories, then three-member categories and finally four-member
categories were encountered.

6.1.2 Data Analysis

The dependent variable in Task 2 was the presence of category names. In other words, we investigated whether the presence of category names made a difference to the informants' performance. The questions were divided into three subgroups according to the numbers of category members they contained. Questions 1-8 contained 2 category members (i.e. 2-member sets), questions 9-12 contained 3 category members (i.e. 3-member sets), and questions 13-18 contained 4 category members (i.e. 4-member sets). The indication of the presence of category names and the numbers of correct answers from the informants for questions of 2-member sets are listed in Table 6.1.

Table 6.1: Presence of category names and numbers of correct answers provided by informants for questions of 2-member sets.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Given			12		12	2	18	
Not given	25	18		15				0

Possible score 28.

The number of correct responses to test items when the category name was present in the co-text (questions 3, 5, 6, and 7) are listed in the row labelled "Given". Similarly, the number of correct responses to test items when the category name was not present in the co-text are listed in the row labelled "Not given", as shown in Table 6.1. The maximum possible score is 28 corresponding to the number of informants involved in the experiment. Thus, question 1 is marked in the second row (Not given), meaning that the question does not contain a category name, and

that the number of correct answers from all the informants is 25 (i.e. the score is 25).

The same listing strategy is used with the other two subgroups, 3-member and

4-member sets.   Tables 6.2 and 6.3 show these results.

Table 6.2:   Presence of category names and numbers of correct answers provided by informants for questions of 3-member sets.

	Q9	Q10	Q11	Q12
Given	21		1	
Not given		19		1

Possible score 28.

Table 6.3:   Presence of category names and numbers of correct answers provided by informants for questions of 4-member sets.

	Q13	Q14	Q15	Q16	Q17	Q18
Given	11		4		8	
Not given		3		3		0

Possible score 28.

As the tables show, the total number of test items where the level of inclusion

is given is equal to the total number of test items where the level of inclusion is not

given.   This is intentional, so that a true comparison can be made between

categories presented with and without category names.   We aggregated and

averaged the scores of those test items containing the same number of category

members with category names present for each informant and contrasted them with

the scores of those test items containing the same number of category members but

without category names present for the same informant. Therefore, for each informant, there were three pairs of scores: the average score for test items containing 2 category members with category names and for the test items containing 2 category members without category names; the average score for test items containing 3 category members with category names and for test items containing 3 category members without category names; and the average score for test items containing 4 category members with category names and for test items containing 4 category members without category names.

Such an arrangement allowed the researcher to discuss the results in relation to Hypothesis 3 pair by pair and identify any difference among these pairs. Descriptive statistics, including means, SD and mean differences for each pair are always shown first in order to give a general impression of the results. For the first pair, viz. items containing 2 category members with or without the presence of category names, the results are shown in Table 6.4.

Table 6.4:   Recognition of membership as a function of presence of category names: means and standard deviations for categories of 2-member sets.

2-member pair	n	Mean	SD	Mean Difference
Given	28	.402	.229	-.143
Not given	28	.544	.205	

In this table, means, SD and mean difference are shown. The mean difference is a negative value, meaning that the scores for the questions with category names present are lower than those questions without category names. This is somewhat surprising because it means that the presence of category names did not enhance the ability to recognize category membership but, on the contrary, seemed to impede the

ability.   To see whether the difference reached a level of significance, we conducted

a paired-samples t-test.   The results are shown in Table 6.5.

Table 6.5:   Paired Samples T-Test for recognition of membership as a function of
presence of category names for categories of 2-member sets.

	SD	df	t	Sig.
2-member pair Given- Not given	.267	27	-2.828	.009

Because of the p value ($p<$ .05), we can see clearly that the difference does

reach a level of significance.   This is a totally unexpected result and exactly the

opposite to what was predicted by Hypothesis 3.   We will return to this finding later

and seek an explanation.   We turn now to test items containing 3 category members.

The means, SD, and mean difference for test items are shown in Table 6.6.

Table 6.6:   Recognition of membership as a function of presence of category names:
means and standard deviations for categories of 3-member sets.

3-member pair	n	Mean	SD	Mean Difference
Given	28	.411	.195	.036
Not given	28	.375	.259	

In this Table, the mean difference shows that the scores for the questions with

category names present are higher than those without category names.   This means

that the informants' performance on recognition of category members was better

when the co-text contained the superordinate category name than when it did not.

This result was the converse of what we found for 2-member pairs. To determine whether the difference was significant, a paired-samples t-test was performed. The results are shown in Table 6.7.

Table 6.7: Paired Samples T-Test for recognition of membership as a function of presence of category names for categories of 3-member sets.

	SD	df	t	Sig.
3-member pair Given- Not given	.302	27	.626	.537

Table 6.7 indicates that the p is larger than 0.05 and hence that the difference is not significant. We cannot therefore conclude that the informants could recognize category members better when presented with superordinate category names in the co-text. Thus, the indication of inclusion did not make a significant difference to the ability to recognize category membership for 3-member sets.

For the last pair, i.e. items containing 4 category members, the same analyzing procedure was repeated, and the means, SD, and mean difference are shown in Table 6.8.

Table 6.8: Recognition of membership as a function of presence of category names: means and standard deviations for categories of 4-member sets.

4-member pair	n	Mean	SD	Mean Difference
Given	28	.274	.288	.202
Not given	28	.071	.139	

As in the case of 3-member pairs, the mean difference is positive, indicating

that the scores for questions presented with category names are higher than those without such labels. That is to say, the informants performed better when the test items were presented with superordinate category names. This result parallels that obtained for 3-member sets and differs that obtained for 2-member sets. In order to determine whether the difference reached a level of significance, a paired-samples t-test was conducted. The results are shown in Table 6.9.

Table 6.9: Paired Samples T-Test for recognition of membership as a function of presence of category names for categories of 4-member sets.

	SD	df	t	Sig.
4-member pair Given- Not given	.246	27	4.357	.000

Interestingly, the result does reach a level of significance and confirms the prediction of Hypothesis 3. Furthermore, the result is the converse of that obtained for 2-member sets. It has been shown that for 2-member sets, successful vocabulary learning defined in terms of the ability to recognize category membership was more likely to occur without the presence of category names. In contrast, for 4-member sets, successful vocabulary learning was more likely to occur with the presence of category names. Before we reach any conclusion and so as to be prudent, in the following section we will look through the test items more closely and check the bracketing patterns of the informants' responses in case it turns out that the means of indicating recognition of membership are problematic. The obvious way of carrying out this check is to look in more detail at items where the results are notably out of line with the broader patterns of responses, as unfortunately a number of the results in this experiment are.

6.1.3 Re-evaluation of Test Items: particular items

Basically, we will look at questions with extreme scores (either much lower or much higher than the norm) and analyze the answering patterns provided by the informants in order to understand why these test items proved difficult or easy for the informants. Categories containing 2, 3 and 4 members will be discussed separately.

6.1.3.1 2-Member Sets

Table 6.1 shows that for 2-member sets, question 6 and question 8 were problematic. Both test items obtained a very low score, 2 for question 6 and 0 for question 8. Question 6 is illustrated below.

Question 6 (2 items)
Strange noises. [Voices in the distance]. [Footsteps going up and down somewhere near, echoing on a stone floor].

(Correct answers are bracketed.)

After checking through the answers given by the informants, we found that six informants bracketed *Strange noises* and *Voices in the distance* as category members and four bracketed *Strange noises* and *Footsteps going up and down somewhere near, echoing on a stone floor* when in fact *Voices in the distance* and *Footsteps going up and down somewhere near, echoing on a stone floor* are category members. That means that the majority of the informants took the category name *Strange noises* to be one of the category members. It is difficult to be sure of the reason for this. Perhaps the context provided in the test item was insufficient or perhaps the superordinate concept of *Strange noises* was more specially defined by the text of

202

the reader than was expectable based on encyclopaedic knowledge. But for whatever reason, the test item failed to invoke the complete context.

Question 8 (2 items)
But [there was something else I had to ask].
'What about Malgosia?' I said, but nobody answered me. 'Come on,' I said. [I was desperate to find out] before I fell asleep again.

(Correct answers are bracketed.)

In the case of question 8, the informants had difficulty recognizing the category structure. Five informants bracketed *there was something else I had to ask* and *'What about Malgosia?'* as category members; one informant bracketed *'What about Malgosia?'* and *I was desperate to find out before I fell asleep again*; the others revealed no particular patterns, just bracketing random phrases. None of the informants seemed to be able to recognize the actual members. This result suggests that the test item is extremely difficult, especially without the presence of a category name.

6.1.3.2   3-Member Sets

Among 3-member items, we paid special attention to questions 11 and 12, both having only one informant who supplied the correct answer. For question 11,

Question 11 (3 items)
We could hear [a conversation], [some shouting]. And then suddenly I heard [a voice I was sure I recognised, a female voice, Malgosia's voice].

(Correct answers are bracketed.)

The majority of wrong answers bracketed *a voice I was sure I recognised, a female voice*, and *Malgosia's voice*. Twelve informants bracketed these two items

as category members. Another four bracketed the former two members but made mistakes with the third, either recognizing *a female voice*, or *Malgosia's voice* as a member. It seems that the informants were misled by the appositive phrases attached to one category member and mistook these phrases for category members. A further look at the structure of this test item showed that these mistakes might result from a degree of rhetorical parallelism among these phrases, which might have been taken to suggest category membership. The influence of rhetorical patterning on the ability to recognize category membership was considered an important factor in designing the methodology and would be addressed in Hypothesis 4.

Question 12 (3 items)
Well this man, Reinaldo, [he noticed your wife]. The red hair, he said, she was beautiful. [He couldn't take his eyes off her]. [She reminded him of an actress he was especially keen on].

(Correct answers are bracketed.)

As shown in Table 6.2, only one informant correctly recognized all three members. Four informants could recognize *He couldn't take his eyes off her* and *She reminded him of an actress he was especially keen on*, and two could recognize *he noticed your wife* and *He couldn't take his eyes off her*. Question 12 involved indirect speech, which might have been one reason for the difficulty the informants had in correctly identifying *he noticed your wife, He couldn't take his eyes off her*, and *She reminded him of an actress he was especially keen on* as category members. This could explain why the informants could typically recognize *He couldn't take his eyes off her* and *She reminded him of an actress he was especially keen on* as category members, but very few could recognize *he noticed your wife*. This might have been caused by the indirect speech that followed.

An alternative explanation to the failure to recognize *he noticed your wife* as a member is that the verb *noticed* is culturally polysemous and thus may have invoked a different reaction in the reader from that intended by the author. That is, the action of noticing may be polite or impolite in different cultures or may result from the relatively attractive in relatively unattractive appearance of the person concerned, and thus the message carried by this item may be congruent or incongruent with the wider text. This may explain why seven informants bracketed *she was beautiful* as a category member, although clearly it does not belong to the three-member set that the informants were intended to identity. Therefore, the performance of the informants could be partly attributed to cultural differences.

6.1.3.3   4-Member Sets

As the number of category members increased, it was inevitable that the sentences became longer and more complicated. Items containing 4 category members thus required consideration with respect to their complexity in terms of sentence structure. Question 13 seemed to be the shortest 4-member item and, perhaps significantly, more informants were able to recognize all four members. In contrast, questions 16, 17, and 18 were more complex sentences. Very few informants were able to recognize the category members in question 16 though these members showed a degree of semantic equivalence as in the following illustration.

Question 16 (4 items)
[This time I was determined to open my eyes].
[I managed to raise the corner of one eyelid] with what seemed like a great effort. I was blinded by bright white light and shut it again. There was someone by the side of my bed. My bed? I was in a bed? What on earth was going on? [I forced my eye open again], and [managed to keep it open for a second longer].

(Correct answers are bracketed.)

205

The four category members, i.e. *This time I was determined to open my eyes, I managed to raise the corner of one eyelid, I forced my eye open again, managed to keep it open for a second longer*, were all indicating an determination of the character to keep awake in the hospital bed after serious injuries. On the contrary, a few informants were able to identify all category members correctly in question 17 perhaps because the cultural knowledge reflected in it, i.e. the schema of a hospital, was more or less the same as that of the informants' cultural background.

In addition, question 18 was especially problematic because, unlike items 16 and 17 where encyclopaedic knowledge could be appealed to. In this item,

Question 18 (4 items)
'Malgosia!' I cried, 'Malgosia!' and ran down the hillside. They'd heard me now. [Tibor turned round and looked to see where the noise was coming from]. [The other man stopped]. [Two more men ran out on to the terrace and looked up]. [They pulled out guns]. 'Malgosia!' I cried again.

(Correct answers are bracketed.)

the category members were describing a series of actions of the scumbags when they found that someone might be overhearing their illegal plan. Without close understanding of the text about the causes of the event, it was difficult for the informants to recognize the category structure and its members.

The purpose of examining individual items in this section was not to justify the adequacy or inadequacy of each test item in a *post hoc* way, but to show how assessment of vocabulary knowledge, especially as defined in this research, eventually involves the complexity of the containing schema and the cultural knowledge of the informants, and that to measure this knowledge requires

consideration for these aspects and thus cannot be achieved until the first attempt is made. This discussion may provide guidelines in selecting test items for future research.

6.1.4 Discussion

Even though the explanations discussed above are those of the researcher, they are based on the bracketing patterns provided by the informants. The reversed results for 2-member and 4-member items, both reaching a level of significance, draw attention to another factor which might contribute to the apparently contradictory findings, namely the number of category members present in a stretch of text. Literature was cited in the methodology chapter which showed that the number of exemplars could make a significant difference in category abstraction. Gentner & Namy (1999), for example, demonstrated that providing children with multiple instances within a category was more likely to enable them to choose taxonomic alternatives, i.e. objects belonging to the same taxonomic category but perceptually distinct from the instances previously shown to them. But, when provided with single instances, children chose perceptually similar alternatives, i.e. objects with similar shapes but belonging to different categories, and taxonomically related ones at equal chance. They concluded that multiple instances invited the children to notice conceptual commonality among category members (Gentner & Namy, 1999).

However, there was an important difference in task design between Gentner & Namy's study and the current study, since Gentner & Namy primed their informants with a standard category member. The task used in the current study did not provide any "standard" category member and therefore would seem to be more

difficult in those cases where category names were present because category members and category names might be confused, with the result that the informants might take the category name as a category member.

As has been indicated already, the presence of category names assisted the recognition of category members in cases where more members in a category were present, but it did not boost this ability when fewer members were present. In the case of items containing two category members, it turned out that recognizing the two category members was not any more difficult when category names were not present in the co-text; indeed contrary to our prediction, the presence of the superordinate category names caused what appeared to be an interference effect so that the informants became confused and were not able to recognize the actual category members as frequently as in cases where the category name was not present. In other words, the level of syntactic complexity of the context in which the test items occurred turned out to be a relevant variable. In this respect, the results of this investigation in terms of the ability to recognize category members differed from those obtained in previous research on categorization.

In addition, the semantic categories present in the text were so particularly defined by the context in which they occurred that without a proper understanding of the overall context, it would not be easy to recognize the category structure. This finding replicated what had been revealed in most of the research studies on the categorization of taxonomic objects. In this context, it should be noted that with the exception of the experimental research in cognitive psychology which manipulates test items in a context-free environment, almost all natural categorization in daily life occurs in a specific context and thus constrains the conceptualisation of the higher-level category so as to reflect the internal structure of

its members. When we think of the category 'apple', what comes to mind in the UK may include 'Cox', 'Gala' or 'Red Delicious', but these will not come to our mind in Taiwan where different varieties exist. It seems obvious therefore that cultural background will contribute to categorization. In a cross-cultural dimension, while language proficiency is held constant, cultural knowledge may have more important effects on categorization. In a word, we want to emphasize the fact that categorization is not totally context-free in practice, especially with respect to instantial or non-conventionalized categories. It is a higher level of context, such as that implicated in cognitive models, which are embodied by our cultural backgrounds to an extent that will be involved.

Finally, although we found reversed results for 2-member and 4-member sets, the investigation of the relationship between the presence of category names and the recognition of category members is still a promising field both in psychology and language learning.

6.2   Experiment 4, Membership and Rhetorical Coupling

In this section, we will discuss the data obtained from Task 3 and the results of statistical analysis in relation to Hypothesis 4. The properties of category membership are still under investigation, but is now associated with another variable that is hypothesized to affect judgments about category membership, degree of rhetorical coupling. As a correlation study, a different kind of statistical instrument is required to analyze the data obtained in this experiment. As in section 6.1.3, special attention will be paid to test items which seem problematic. This does not mean that we favour a *post hoc* exclusion of test items but rather that results which diverge from our expectations and one out of line with the general direction of

results may help us to facilitate future research in the same area.

6.2.1 Hypothesis 4 and Task 3

<u>Hypothesis 4</u>

Successful vocabulary learning (= recognition of category membership) is more likely to occur where multiple instances of membership at the same level in a semantic category share common rhetorical properties.

Hypothesis 4 was designed to investigate the association of the ability to recognize category members and the degree of rhetorical coupling of these members. "Coupling" is used as a cover term to refer to the phonetic or semantic equivalence of items sharing positional or syntactic features. It was predicted that the coupling of semantic denotation and syntactic construction of two or more closely connected items would strongly indicate membership of a category. That is, the higher the degree of coupling displayed by category members, the more likely informants would be to recognize them as members.

Task 3 was designed to test Hypothesis 4. This task adopted a multiple-choice format and required informants to choose a correct answer from among three possible answers. All the three possible answers contained the same complete discourse context in which a set of possible category members was embedded, but differed in the potential category members indicated by means of underlining each of the putative sets of members. The informants were required to identify the answer in which members of a common category were underlined. The task therefore required informants to identify category members correctly.

6.2.2 Data Analysis

Two sets of data will be involved in the analysis. One set was derived from the result of Task 3 performed by second language learners and the other consisted of the ratings of rhetorical coupling provided by native speakers.

Task 3 contained 10 multiple-choice items, each consisting of one correct answer and two incorrect answers. Each correct response was scored 1 to give a maximum score per item of 30 because 30 informants participated in this task. These 10 questions were previously rated by native speakers with respect to degree of rhetorical coupling, with 0 indicating no coupling and 3 indicating the strongest degree of coupling. We calculated the average rating of rhetorical coupling for each question using the following formula.

$$\text{Average rating} = (N_3 \times 3 + N_2 \times 2 + N_1 \times 1 + N_0 \times 0) / 12$$

This formula means that, for each item, the sum of the following was computed:

the number of NS (native speakers) 3 point ratings multiplied by 3

the number of NS 2 point ratings multiplied by 2

the number of NS 1 point rating multiplied by 1

the number of NS 0 point ratings multiplied by 0.

This sum was then divided by 12, 12 being the total number of native speaker raters participating in the research. In this way, an average rating for each question on a scale of 0-3 point was obtained. Table 6.10 shows the result of informant scores and the average rating per question.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
Score	26	18	0	21	23	15	20	14	22	28
Average Rating	0.42	0.67	0.92	2.58	1.17	0.58	1.08	0.92	2.17	2.42

Possible score 30, maximum rating 3.

From this table, we can see that the item 3 is scored 0, and that item 1 obtains a very high score but the rating of degree of rhetorical coupling is very low. Later, we will return to these two items and discuss them in more detail. To see whether there was an association between the ability to recognize category membership and degree of rhetorical coupling, we performed a Pearson correlation test on these two sets of data. The result showed that the correlation coefficient was low ($r= 0.370$, $p> .05$), and hence that there was no strong correlation between the scores and average ratings.

In addition, we made an attempt to rank each question within each set of data. The results are shown in Tables 6.11 and 6.12.

Table 6.11:   Rank order of scores for all questions.

Rank	1	2	3	4	5	6	7	8	9	10
Question	Q10 (28)	Q1 (26)	Q5 (23)	Q9 (22)	Q4 (21)	Q7 (20)	Q2 (18)	Q6 (15)	Q8 (14)	Q3 (0)

Table 6.12:   Rank order of average ratings for all questions.

Rank	1	2	3	4	5	6	6	8	9	10
Question	Q4 (2.58)	Q10 (2.42)	Q9 (2.17)	Q5 (1.17)	Q7 (1.08)	Q3 (0.92)	Q8 (0.92)	Q2 (0.67)	Q6 (0.58)	Q1 (0.42)

In each table, the parenthesized figure shows score (Table 6.11) and average rating (Table 6.12) for each question.   Notice that in Table 6.12, question 3 and question 8 have the same average ratings and are therefore ranked at the same position.   Based on these rankings, a Spearman test was conducted.   This showed that the correlation coefficient was again low ($r = 0.365, p > .05$).

At this stage, we could not find any association between the ability to recognize category membership and the common rhetorical properties shared by category members, therefore disconfirming Hypothesis 4.

6.2.3   Re-evaluation of Test Items: particular items

In this section, we will look in more detail at question 3 where no correct responses were recorded and at question 1 where the second highest number of correct responses was recorded despite the item having the lowest rated degree of rhetorical coupling.

It was strange that none of the informants was able to give the correct answer for question 3, especially when all the other questions obtained approximately half the possible score or more.   This is the item:

Question 3
a) <u>It was a bumpy journey, with turbulence almost all the way.</u> Most of the time I half slept as the plane bounced through the skies.
b) <u>It was a bumpy journey,</u> with turbulence almost all the way. Most of the time I half slept as <u>the plane bounced through the skies.</u>

c) It was a bumpy journey, <u>with turbulence almost all the way</u>. Most of the time I half slept as <u>the plane bounced through the skies.</u>

<p align="right">(b is the correct answer.)</p>

The data shows that out of 30 informants, 18 chose answer a) and 12 chose c). Answer a) grouped *It was a bumpy journey* and *with turbulence almost all the way* as category members, and answer c) grouped *with turbulence almost all the way* and *the plane bounced through the skies*, as category members. The correct answer was b) in which *It was a bumpy journey* and *the plane bounced through the skies* were grouped as category members. We speculate that it was the degree of implicit causal linking between the wrongly supposed category members in options (a) and (c) that led the informants to take them as related.

The results for question 1 strongly contradicted our hypothesis in that the category members displayed a very low degree of rhetorical coupling as indicated in the NS ratings, but the informants readily identified them.

Question 1
a) <u>Her hair was cut short</u>. She wore heavy black glasses. She watched me as I struggled in with my suitcase. <u>She didn't look welcoming at all</u>.
b) Her hair was cut short. She wore heavy black glasses. <u>She watched me as I struggled in with my suitcase</u>. <u>She didn't look welcoming at all</u>.
c) <u>Her hair was cut short</u>. She wore heavy black glasses. <u>She watched me as I struggled in with my suitcase</u>. She didn't look welcoming at all.

<p align="right">(b is the correct answer.)</p>

A substantial number of informants, 26 out of 30, chose the correct answer. We speculate that the two verbs in the two members sentence, i.e. "watch" and "look", suggested membership to the informants because of their semantic similarity. However, if we temporarily disregard this item and conduct a Spearman test on the rankings of both sets of data, excluding question 1, we obtain a very high correlation

<p align="right">214</p>

($r = 0.745$) significant at ($p < .05$). This means that Hypothesis 4 is confirmed and that there is a strong correlation between the ability to recognize category members and the degree of rhetorical coupling of these members. However, we do not feel justified in excluding question 1 and in treating it as exceptional only in order to support our hypothesis. We therefore suggest a more modest conclusion, that, due to the small number of test items, the results may not be fully reliable and that a large-scale study of this kind is necessary to determine whether the hypothesis is confirmed or falsified (although we would expect it to be confirmed).

6.2.4 Discussion

One important reason why we were not able to obtain a definite result to either confirm or disconfirm our hypothesis is that the number of the test items, 10 in total, is too few to establish a correlation. Even if we obtained a high correlation value between these two variables, i.e. the ability to recognize category membership and rhetorical patterning displayed by category members, we would still need to be careful in making a strong assertion about their association. Ideally, at least 30 items are required for a valid correlation test (Charles, 1995: 198; McMillan & Schumacher, 1997: 176). Nevertheless, to identify 30 suitable test items would require at least 3 chapters of the reader, a situation which would demand an unrealistic quantity of text to be read by the second language learner informants in relation to a single pedagogic test. Obviously, the present research was constrained in such a way that only a small-scale experiment was practical in the normal classroom setting. One may alternatively suggest obtaining data on three different occasions so as to have sufficient amount of data for correlation tests. However, this was impossible in the context of the present experiment because a variety of task

types rather than a single repeated type was required for pedagogical reasons.

This experiment may therefore be treated as a pilot study and the hypothesis can be seen as very promising, especially if we disregard the two problematic items (i.e. questions 1 & 3). We should also bear in mind one of suspicions that occurred in an analysis of Experiment 3, that rhetorical patterning between members does suggest category membership. It is not only the possibility of confirming our hypothesis that intrigues us, but also, given the disparity between the result for question 1 and for all the other questions, the possibility that second language learners and native speakers do not share the same intuition with regard to degree of rhetorical coupling.

6.3 Experiment 5, Degrees of Membership and Rhetorical Coupling

This section will focus on another aspect of category membership, the degree of membership (or prototypicality) of category members. In contrast to the prediction of Hypothesis 4, Hypothesis 5 expects a reverse association between recognition of degree of membership and degree of rhetorical coupling. This is because distinct rhetorical properties perceived among category members are assumed to suggest a finer differentiation of degrees of category membership and because rhetorical coupling is typically taken to mask perception of meaning differences. Similarly, analysis of the results will involve correlation tests on two different sets of data representing two different variables. As in explaining the results for Experiment 4, we will again be cautious about making strong claims given the limited database.

6.3.1 Hypothesis 5 and Task 4

<u>Hypothesis 5</u>

Successful vocabulary learning (= recognition of degree of category membership) is more likely to occur where multiple instances of membership at the same level in a semantic category have distinct rhetorical properties.

Hypothesis 5 was designed to test whether rhetorical coupling would be negatively associated with the ability to recognize the degree of category membership. It was predicted that the degree of prototypicality attained by category members would be more readily recognized in cases where the degree of rhetorical coupling was lower. This is because rhetorical coupling is associated in the literature with drawing attention to message form at the expense of message meaning (Jakobson, 1960; Mukarovsky, 1970). In other words, if the lexical sequences used to denote category members exhibited relatively distinct rhetorical properties, a finer distinction of degree of category membership would be more apparent.

Task 4 was used to test Hypothesis 5. Each test item in this task consisted of a superordinate category name and a set of members, with four possible degrees of membership provided for each answer, **a very good, a good, an acceptable,** or **not a good** member. Informants were asked to rate the prototypicality of each member within a category on this four-point scale. This format allowed the researcher to determine whether informants were able to differentiate each category member with respect to degree of membership of the given category.

6.3.2 Data Analysis

As in Experiment 4, the data analysis involved two sets of data, one provided by the informants and the second consisting of ratings of rhetorical coupling provided by native speakers.

Since the informants might draw on their background cultural knowledge in making judgements and since it was likely that in each item there was an inherent degree of prototypicality for each member (see p. 146), we distinguished only between cases where any degree of membership difference was identified (scored as 1)[1] and cases where no degree of membership difference was identified (scored as 0).

Table 6.13:   Results of scoring: scores, and average ratings for all questions.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
Score	22	30	21	26	22	30	21	19	23	21	28
Average Rating	1.92	1.50	0.83	1.42	2.00	1.50	0.50	1.42	1.50	0.83	2.17

Possible score 30.

In this table, the average ratings given by native speakers are also shown. The calculation of average ratings is described in section 6.2.2. To determine whether there was a correlation between these two sets of data, we performed a Pearson correlation test on them. The result showed that, contrary to our prediction, there was a positive correlation coefficient, rather than a negative one ($r = 0.394$, $p > .05$).

---

[1] A response was scored 1 whether an informant identified a degree of membership difference across two levels (e.g. from **a very good** to **a good**) or across three levels (e.g. from **a very good** to **an acceptable**) or across all four levels (e.g. from **a very good** to **not a good**).

According to Hypothesis 5, it was assumed that the more distinct the rhetorical properties of category members, the more readily the degree of category membership would be discriminated. In statistical terms, a negative correlation coefficient would be expected. That is, no evidence could be found to support the hypothesis. In order to perform a Spearman correlation test, we then ranked each question within each set of data, as shown in Tables 6.14 and 6.15.

Table 6.14:   Rank order of scored questions.

Rank	1	1	3	4	5	6	6	8	8	8	11
Question	Q2 (30)	Q6 (30)	Q11 (28)	Q4 (26)	Q9 (23)	Q1 (22)	Q5 (22)	Q3 (21)	Q7 (21)	Q10 (21)	Q8 (19)

Table 6.15:   Rank order of average ratings for all questions.

Rank	1	2	3	4	4	4	7	7	9	9	11
Question	Q11 (2.17)	Q5 (2.00)	Q1 (1.92)	Q2 (1.50)	Q6 (1.50)	Q9 (1.50)	Q4 (1.42)	Q8 (1.42)	Q3 (0.83)	Q10 (0.83)	Q7 (0.50)

In both tables, the parenthesized figure indicates score (Table 6.14) and average rating (Table 6.15) for each question. Also notice that in Table 6.14, questions 2 and 6, questions 1 and 5, and questions 3, 7, and 10 are ranked at the same positions because they have the same scores. In Table 6.15, questions 2, 6, and 9, questions 4 and 8, and questions 3 and 10 are also ranked at the same point because in each of the three sets, the rating for the individual questions are the same. Based on the rankings, a Spearman correlation test was conducted and the coefficient was a positive value ($r = 0.579$, $p > .05$). This result contradicted the prediction of Hypothesis 5, which expected a negative correlation.

Given that the result did not confirm our prediction, we tried to consider other possibilities. That is, since the numbers of category members differed from question to question, the scoring might be influenced to some extent by this variable, bearing in mind the effect of the number of category members on recognition of category membership found in Experiment 3. It might reasonably be expected that an item with more category members would be scored higher than one with fewer members since there might be a greater likelihood of degrees on membership being found in a larger set of members. To check this possibility, we constructed another table showing the numbers of category members and scores for all questions:

Table 6.16:   Numbers of category members and scores for all questions.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11
No. of Members	2	4	2	3	2	5	2	2	2	2	3
Score	22	30	21	26	22	30	21	19	23	21	28

As shown in Table 6.16, the questions with more category members do score slightly higher than those with fewer members. Therefore, we conducted a Pearson correlation test again only on those questions with an equal number of category members to exclude the influence of this factor. Only questions containing 2 category members were taken into consideration, i.e. questions 1, 3, 5, 7, 8, 9, and 10, because there were more test items in this group. Similarly, we did not find a negative correlation as required to confirm the hypothesis ($r= 0.347, p> .05$). Even when we conducted a Spearman correlation test, the result was the same ($r= 0.623, p> .05$).

As can be seen, the hypothesis could not be confirmed though we tried to take out a possible variable, i.e. the number of category members. We therefore had to accept that distinct rhetorical properties among category members are not associated with more successful vocabulary learning when defined in terms of the ability to differentiate degree of membership.

6.3.3 Discussion

In structuralist poetics, it is widely accepted that there is a trade-off between degree of rhetorical patterning and meaning processing. That is, when the degree of rhetorical patterning is higher, meaning will be less salient and less processable; and *vice versa*. Since graded readers are a form of literature, we would expect this trade-off to apply to them too. However, we could not find any support for this assumption, based on the data collected and analyzed to test Hypothesis 5. We are therefore obliged to conclude that differentiation of degree of category membership is a robust ability which is not affected by the superficial phonological patterns or other rhetorical features shared by the category members themselves.

One possibility is that the readers, i.e. the informants, adopted a strategy of bottom-up processing, presumably because they were not so advanced as we had expected. While the understanding of rhetorical properties required the ability to grasp the suprasegmental patterning, it seemed difficult for this group of readers. That is, the informants might not be able to detect the diversity of rhetorical properties as presented in category members and thus unable to differentiate the degree of category membership. This implies that the processing strategies adopted by the learner in reading can affect the overall understanding of the text.

This study also has practical implications with respect to the association of

rhetorical coupling and language teaching. Two bipolar aspects of rhetorical coupling were investigated in this study: the commonality and the diversity of the rhetorical properties of lexical phrases. The former was assumed to enhance the ability to recognize category membership in that the phonetic/semantic and structural/syntactic equivalence might be taken as kind of similarity between category members and thus highly suggestive of common category membership. In contrast, it was supposed that rhetorical diversity, would allow the meaning of category members to be better processed and enable the informants to differentiate degrees of category membership. Although it is widely acknowledged that when the language form is more poetic, the message content is less salient, we found no support for this argument based on the results of this study.

6.4 Conclusion

In this chapter, an additional property of semantic categories, category membership, was investigated. Category membership may be considered fundamental in acquiring vocabulary depth knowledge and involves both recognition of membership and discrimination of degree of membership. As the experimental psychology literature reveals, the presence of category names leads to more successful recognition of category members in infants (Markman, 1987). To investigate the extent to which this would be replicated in second language learning, informants were presented with pairs containing the same numbers of category members with and without category names. It was found that in the case of 2-member pairs, the presence of category names did not lead to more successful vocabulary learning, whereas with 3- and 4-member pairs, the presence of category names did seem to lead to more successful recognition of category membership.

This finding differs slightly from the findings of most psychological studies of infants and categorization behaviour and suggests that the particular linguistic nature of the context affects the recognition process.

In Experiments 4 and 5, a further variable was introduced which was hypothesized to influence the informants' ability to recognize both category membership and degree of category membership. By involving native speakers as raters of the degree of rhetorical patterning displayed by category members, the researcher was able to determine whether there was an association between degree of rhetorical coupling and recognition of category membership. In Experiment 4, no significant result was found when the pre-determined methodology was employed, but an alternative analysis which had not been foreseen when the methodology for this study was devised justified the underlying hypothesis. In other words, the assumption underlying the hypothesis still looks promising.

Whilst a high degree of rhetorical patterning was assumed to suggest common category membership, distinct rhetorical properties in sets of category members was assumed to lead to enhanced differentiation of degree of category membership. In Experiment 5, an attempt to prove an association between a low degree of rhetorical patterning among category members and discrimination of degree of category membership failed in that the correlation test did not confirm our hypothesis. Because this result was so unexpected, we tried an alternative way of making sense of the data in order to take into account an unexpected variable possibly influencing the original scoring, i.e. the number of category members. Nevertheless, the results were far from affirmative. We therefore had to admit that the results disconfirmed the hypothesis and we thus concluded that the differentiation of degree of category membership is highly cognition-based and will not be affected by the superficial

rhetorical features of category members.

Finally, one important limitation of this study needs to be admitted. The small number of test items involved in the investigation of Hypotheses 4 and 5 makes us cautious about making any strong claims about the relationship of rhetorical coupling and the ability to recognize either category membership or degrees of category membership. This basically stems from the insufficient number of test items for performing correlation tests and thus makes the results not so reliable as they should be, for which shortcoming there were practical pedagogic reasons in this case.

Chapter 7 Implications

In this chapter, we will discuss the implications of the findings obtained from the experiments described in chapters 5 and 6. In the first section, the limitations of the research will be admitted, given the pedagogical constraints inherent in classroom-based research and its effect on the quality and quantity of the data obtained. Section two will be concerned with language pedagogy, especially with respect to vocabulary depth, and will discuss the practical implications for the association of vocabulary depth knowledge and reading comprehension investigated in this thesis. Section two also discusses the development of reading materials. Section three will discuss categorization behaviours revealed in this study. The results obtained from the instantial categorization investigated here will be considered and contrasted with those found in experimental psychology. The discussion will focus on the effects of the presence of category names and of the number of category members on the ability to recognize category membership. Section four will provide a brief summary of all the implications that have been discussed, whilst the last section will reveal the significance of this study, especially in relation to vocabulary learning from context and the methodology used in this research.

7.1 Limitations of the Research

Before discussing the implications of this research, we have to recognize the contextual constraints and limitations inherent in this study. First of all, only a single proficiency level was chosen. In chapter 4 we justified working with advanced level learners since it was vocabulary depth knowledge that was undergoing investigation. Vocabulary depth knowledge was defined as the ability to recognize the inclusion and

membership properties of semantic categories in a natural linguistic context consisting of an extensive text. The role of the linguistic context was assumed to facilitate learning of vocabulary depth knowledge when understood in terms of inclusion and membership. We return to this limitation in section 7.2.4. because it might provide an opportunity for future research.

Secondly, only the work of a single author was covered in this research. The author selected for this research was an experienced language teacher and a well-known materials writer across a wide range of fields including graded readers, coursebooks for learners and methodology texts for teachers. His experience as a successful language teaching materials writer suggested that his work would be likely to achieve successful transmission of meaning for second language learners. Since the time available to collect data within the English course was limited, this constituted a further constraint on the number of graded readers that could be included, so that effectively we were limited to working with a single reader.

Another significant obstacle stemmed from the time allocated for the work based on the graded reader. The peripheral status of readers within the wider curriculum was commonplace, and applied also to the course in the university in which this research was conducted. Accordingly, we were not allocated as much time for work on the text we were investigating as would have been ideal. This is partly because we took the methodological decision to limit the experimental nature of this research so as to reflect the teacher's style and the schedule of the institution. We also relied on the teacher's skill and judgment in administering the reading comprehension tasks in the classroom and ensuring that every informant read the chapters assigned for pre-class outside reading. The tasks or tests were conducted as a part of the regular teaching

routine and contained various types of activity suitable for pedagogic purposes and already relatively familiar to the learners.

In consequence of the limited time available for the experiments, as described above, the amounts of data that could be obtained were relatively small. This drawback did not affect Experiments 1 and 2 but it did influence the reliability of Experiments 3, 4, and 5. Especially, experiments 4 and 5 required at least 30 test items to assure reliability of the correlation analysis which was conducted. Apart from the time available for this research, the reading speed of the informants also limited the numbers of test items that could be accommodated within the pedagogic schedule. Nevertheless, analysis of the data revealed valuable and useful information for vocabulary learning in terms of the ability of learners to recognize the inclusion and membership properties of the semantic categories present in an extended text. The design of the experiment also reflected the fact that nobody has done this kind of research before and that consequently we had to measure vocabulary depth knowledge with original instruments, which might prove useful in the future. Nevertheless, we have to be cautious about whether our methodology achieved the levels of reliability and validity required in experimental research.

7.2 Implications for Vocabulary Depth Knowledge

Vocabulary depth as defined in this study is concerned with the ability to recognize the internal structure of semantic categories present in an extended text. This definition provides us with an opportunity to investigate the association of vocabulary depth, one of the commonly studied areas in second language vocabulary learning, and studies of reading comprehension. Moreover, implications can be drawn in relation to the development of reading materials as a language input type, especially

with respect to vocabulary learning and teaching. Based on the results obtained from our experiments, the following discussion will thus be concerned with what such a definition entails with respect to reading comprehension, development of reading materials and vocabulary teaching and learning.

7.2.1 Reading Comprehension

Reading is important for language learning because it provides the learner with the opportunity to experience a language and thus gradually consolidate and expand linguistic knowledge in relation to that language. As discussed in chapter 2, reading for comprehension and reading for vocabulary learning are sometimes regarded as competing activities, although in this research they are not in conflict. Indeed, one merit of this study is that it shows that reading comprehension and vocabulary learning can be interdependent in that vocabulary knowledge can only be obtained by understanding the relations between the lexical items embedded in a text.

To be more specific, the experiments conducted in this research are intended to reveal the mechanisms that are employed in understanding a text as a cognitive function. This study explores the likelihood that conceptual categorization could be used to understand the text by reconstructing the semantic categories through which the author's meaning is conveyed. Rather than the object categories usually investigated in experimental psychology, most of the categories investigated in this study are related to the events, actions and characters occurring in the fictional text. Given that it is a general practice that an author creates text-specific superordinate categories all the time, categorization is therefore taken as one of the crucial mechanisms that readers can appeal to in order to understand an extended text.

More importantly, what distinguishes this study from others is the hierarchical relations (inclusion) and internal category structure (membership) of the lexical items and phrases present in the text. Reading for vocabulary learning in other research has emphasized the learning of discrete items as though they are unrelated. Instead, this study assumes a cognitively motivated relationship inherent in these to-be-learned items. It is this characteristic that rendered reading comprehension and vocabulary learning complementary.

7.2.2 Development of Reading Materials

As the results showed, when the superordinate category name and its members were modified so as to be structured more paratactically than the original text, there was a difference in the informants' ability to recognize the hierarchical relationship between category members and a superordinate category name. This finding raises the issue of how a writer presents a text to readers so as to achieve maximum comprehension via categorization. That is, the way an author arranges ideas with respect to structuration may affect the readability of the text. Although many factors are found to influence the readability of a text, including its linguistic complexity, the skill of a reader and motivational factors (Gilliland, 1972), vocabulary difficulty is characteristically regarded as the best predictor of readability (Nation & Coady, 1988: 97). While the frequency of words is widely acknowledged as an important indicator of the ease with which the reader understands a text, this study shows that recognition of the inclusion and membership properties can also be regarded as an indicator of successful comprehension, and thus of readability.

One contribution of this research is in providing additional guidelines for writers and publishers of graded readers. For writers, one way to achieve more effective

transmission of meaning appears to lie in the arrangement or encoding of a superordinate category name and its members. For publishers, although they recognize the influence of vocabulary difficulty on text readability and thus set limits on the vocabulary used at each level of reader, they do not yet acknowledge the role of vocabulary depth in affecting text readability, especially with regard to the construction of categories. This is largely because of the difficulty of knowing how to do this, a problem to which this research perhaps contributes at least a partial solution. This research therefore suggests a new dimension for checking the readability of graded readers.

Furthermore, graded readers are always graded in accordance with the degree of lexical and syntactic difficulty of the text. An insight resulting from this research is that a concern with vocabulary depth as defined in this study can be taken as a criterion for proficiency level. Thus, there is a critical proficiency level at which the internal structure of a category becomes comprehensible to learners. Beyond this level, the inclusion and membership properties of semantic categories can therefore be integrated into a text according to different levels. Needless to say, how to achieve this appropriately for each level is an empirical issue that remains to be investigated.

7.2.3 Vocabulary Teaching and Learning

Traditionally, there are thought to be two ways of learning vocabulary. One is by reference to objects in the real world, the basic strategy of the earlier stages of both first and second language learning. The other is by reference to words in other languages, i.e. by translation equivalents. This is the method frequently practised in second/foreign language learning where vocabulary knowledge is acquired via the mediation of the native language. Each of these methods privileges breadth (extent of

230

vocabulary knowledge) over depth (how words are used). Apart from these two acknowledged methods, this research investigates a third way of vocabulary learning; that is, by reference to other contextualized lexical items. To learn about a lexical item means, amongst other things, learning about the internal structure of the category of which it is a member. By appealing to the internal structure of categories within a specific context, one is able to acquire vocabulary depth with respect to the categorical properties of the lexical items within a text.

As a matter of fact, in many established language teaching methods, a lot of language teaching is merely a matter of making connections outside the language being learned, either with objects presented as realia or in the form of visuals, or with a mother tongue lexical equivalents (translating). Such approaches ignore our natural or intuitive understanding of the way the world is organized. As revealed in this research, to learn a lexical item embedded in a text is to reconstruct the categories created by a writer. Recognition of the inclusion and membership properties of items contained in a text is assumed to lead to more successful learning of these items. This study suggests that language teaching can be more effective if it involves this kind of practice.

This investigation also provides an additional method for studying vocabulary knowledge. We have talked about vocabulary knowledge as involving aspects of breadth and depth, and vocabulary learning and teaching therefore have to take account of these two aspects. As revealed in this research, one way of looking at depth knowledge is taking a context as a superordinate notion, so that the superordinate category name is the meaning of the members of the category that are included within it. The structure of the category defines the meaning of the members in terms of their inclusion and membership properties. Accordingly, the ability to recognize these

properties is considered to be a manifestation of vocabulary depth knowledge. In contradistinction to other theories of vocabulary depth knowledge, the meaning of a lexical item cannot be isolated from its context. Rather, the meaning of a lexical item is defined in its relation to other lexical items found in the context in which it occurs. From this, it follows that the internal structure of categories varies to some degree with the context in which these categories occur, a finding confirmed independently in research described in chapter 3.

7.2.4 Future Work

In this research, the experiments were conducted on second language learners classified as advanced in terms of proficiency because it was assumed that understanding vocabulary depth as defined in this study would require more advanced linguistic knowledge. As a consequence, proficiency level was not treated as a variable. In future, it is recommended that this factor should be included in the experimental design to test for any effects on the understanding of categorization occurring in a text.

Furthermore, due to the time constraints that applied in the experimental setting, only one graded reader could be covered for this research. This also means that only one genre and one authorial style were investigated. Future work can expand the scope of investigation by selecting a variety of genres and include the work of more than one author. The effects of categorization on reading comprehension may well turn out to vary with different texts and text genres and the different ways of organizing superordinate categories appropriate to them.

7.3 Implications for Categorization Theory

Since categorization is the essential mechanism employed in this research, which explores the association of inclusion and categorization with language learning, it is also possible that the categorization patterns or behaviours revealed in this research can contribute to categorization theory to some extent. Basically, there are discrepancies between what had been found in the experimental psychology and the findings revealed in this study. In this section these differences are discussed and their contribution to categorization theories considered.

7.3.1 the Presence of Category Names and the Number of Category Members

This research is valuable because it has revealed two effects which are not predicted by existing psychological research studies. First, the presence of category names/labels was supposed to enhance the ability to recognize category members, but it was shown in this study that the presence of a category name did not aid the recognition of category members for categories consisting of 2-member sets. On the contrary, the presence of a category name or label resulted in a situation in which the informants frequently took the category name or label as one of the category members. The results for 2-member sets therefore contradicted the prediction already established in experimental psychology. This suggests that we need to think carefully about semantic categories in their linguistic context. Recognition of instantial cases may not follow the presumption of psychological research studies concerning categorization, where the distinction between category names, often used as primes, and category members is taken for granted in a way that this research throws some doubt on.

Second, in the experimental literature, the presence of a larger number of category members has been shown to be an influential factor in helping informants to recognize the structure of a category. Although this study did not particularly look into the effect of this factor, it did reveal that the presence of a larger number of category members did not assist category abstraction. This was presumably due to the complexity of the syntactic structure involved and the length of sentences required. Once again, real world linguistic context is found to complicate the categorization task. And once again, the results obtained in these experiments with a natural language text do not replicate those reported in the experimental literature.

Given these results, we can understand that the categorization of events and actions found in a text is not simple. Many factors appear to be involved. First, there is normally no single-word label for the events, actions and personality categories investigated in this research. A string of words presented as a category name or label may activate different reactions to any aspect of any word in the multi-word string. This is not comparable to single-word object categories which seem to have more definite activation paths in our semantic memory. Therefore, the categorization patterns actually found in this study cannot completely accord with the findings derived from the studies of object categories, and especially those where simple category names or labels occur and are used as primes. In this study, the category name itself has to be identified and perhaps in a way that suggests that what are commonly treated as primes in experimental psychology, i.e. category names, are precisely what has to be identified. Put another way, for the learner or reader it may well be that the "primes" are category members rather than category names.

Alternatively, we may attribute the difference of categorization to the means by which categorization is established. In the literature, researchers have demonstrated

the effects of different modes of presentation on categorization behaviours. In deciding whether two exemplars belong to the same category or whether an item is a member of a certain category, the decision time is faster for pictorial than for verbal presentations (Rosch, 1975; Pellegrino, Rosinski, Chiesi, & Siegal, 1977; Hogaboam & Pellegrino, 1978). The verbal statements used in this research and recovered as text might also affect the ease of recognition of category structure or categorization behaviour. Alongside Barsalou's study (1983), this study is one of relatively few studies that focus on non-object, real-life categories of events and actions presented in the form of verbal statements.

Second, most of the categories of events, actions or personality were as a matter of fact shaped by the causal relations present in the text. That is to say, their dependency on context was much higher than would be the case in experimental psychology in that they required a considerable quantity of information to determine their references and interpret their meanings. While the internal structure of a few categories, such as "city" or "hospital", could be largely found in our world knowledge, others were highly constrained by the text, such as the category derived from the situation, "I am not cut out for heroism or dramatic gestures". This implies that we need to have a better understanding of the function of categorization. We need to take into consideration the context factor and put more emphasis on non-object categories rather than on the natural and nominal kinds studies in the experimental literature. This perspective also draws attention to the flexibility of our conceptual organization and its consequent variability across contexts.

Moreover, we speculate that the concepts activated by the category names or labels must have invoked personal experience of prior categorization judgments and, as a result, might have affected the recognition of both inclusion and membership

properties as presented in the context currently encountered. Thus, the label "the life of the river" could well have reminded the informants of something along a creek in their hometown, or of something they read about the Yangtze River in China as a part of a geography course. This means that personal experience and encyclopedic knowledge also have a role in enabling or disabling what in this study has been called instantial categorization. This is certainly an area of rich promise for future research, especially in a world in which global intercultural communication is becoming a regular behaviour for large numbers of people whose particular cultural experiences differ.

Finally, it is necessary to take a broader view of categorization theory in light of the current study. A major difference lies in the category types studied by Rosch (1973b, 1975) and those investigated in this research. Whilst conventionalized categories were used in Rosch's studies, this research investigated instantial or non-conventionalized categories. Conventionalized categories such *fruit* or *furniture* are well-established in our conceptual system, whereas non-conventionalized categories are highly context-dependent. Although seemingly different, the theory of ICMs (Idealized Cognitive Models) is capable of accounting for both. ICMs are the mental schema that contain these two types of categories, and themselves are constructed out of our world experience. This explains why cultural or encyclopedic world knowledge can affect categorization behaviours regarding to these two types of categories, and, as found in this research, instantial or non-conventionalized categories can be heavily influenced by cultural or encyclopedic world knowledge. A cross-cultural investigation of these two types of categories might demonstrate more clearly how these knowledge types are related to categorization.

7.3.2 Future Work

Owing to the limited quantity of data available in a normal class setting, the results of the experiments presented in this study could not always enable us to reach a definite conclusion. Especially for experiments 3, 4 and 5, the number of test items proved too small to produce findings in which we could be confident. More data is necessary in order to obtain a reliable conclusion. However, this is not to say that classroom-based research where only a small quantity of data can be obtained is not worth pursuing, only that the actual nature of language learning involves so many variables that it is far from easy to control them and to measure their effects.

7.4 Summary of Implications

In this chapter, a number of implications are identified from the experimental findings with respect to both vocabulary depth learning in a second language and cognitive categorization. The pedagogic implications include a new way of looking into the association of vocabulary learning and reading comprehension, an alternative criterion for the development of reading materials, and a rationale for vocabulary learning and teaching that could be informed by a new definition of vocabulary depth. To be more specific, this study suggests a new definition of vocabulary depth learning in terms of the inclusion and membership properties of the categories present in a text. This study also associates reading comprehension and vocabulary depth learning in a way that can constitute a new workable principle for the development of reading materials alongside the existing breadth approach. Moreover, the definition of vocabulary depth knowledge proposed here has an impact on the methodology of vocabulary teaching and learning, and provides a complementary strategy to the existing practice of making reference to real objects or to translation equivalents,

thereby enabling both breadth and depth knowledge to be taught. The new way of teaching vocabulary depth knowledge recommended here is more natural and intuitive in that the psychology reality of cognitively-based categorization can be applied in language learning.

With regard to categorization, since the results obtained in this research differ from those reported in the experimental psychology literature, we need to consider the likely effects of the variables particular to the pedagogic context. First, since the presence of category names was not found to have a conclusive effect on the ability to recognize category membership, we speculate that, since the nature of event categories investigated in this research was instantial, the results were different from those obtained from the investigation of object categories in experimental psychology research. Second, without any affirmative findings in relation to the effect of rhetorical coupling either on the ability to recognize category membership or on the ability to differentiate degree of category membership, no definite conclusion could be reached. The categorization theory with respect to instantial categories or non-conventionalized categories requires more investigation.

7.5 Conclusion

7.5.1 Significance of this Research

The significance of this research lies in the new definition of vocabulary depth knowledge it proposes. Since vocabulary depth is defined in terms of the ability to recognize the internal structure of semantic categories occurring in a text, vocabulary depth learning is proved to be closely related to the understanding of the text. Without a proper understanding of the message conveyed in the text, the categories cannot be identified and thus vocabulary depth learning will not be successful. This empirical

finding fills a gap in the second language vocabulary learning with respect to characterizing vocabulary depth knowledge.

In addition, as discussed in chapter 2, although context is agreed to be conducive to vocabulary learning, there is no existing research that demonstrates a notion of context that can be operated experimentally. An appeal to the internal structure of semantic categories present in a text can provide a principled definition of context so that the context of a text is characterized in terms of the semantic categories constructed by an author, and the structure of these categories is manifested in the properties of inclusion and membership. The results of the research prove that a more cognitively oriented mechanism, categorization, can be employed to facilitate language learning.

7.5.2 Methodological Concerns for Future Work

This research set out to investigate the relationship between vocabulary depth knowledge as redefined in this study and semantic categorization. Because this is an entirely original approach to vocabulary learning, it was necessary to construct a largely original method to test the informants' target knowledge, i.e. vocabulary depth knowledge. Therefore, it is not only the nature of the informants' target knowledge but also the methodology employed in this research that should undergo careful scrutiny. Specifically, we need to reflect on the way we defined vocabulary knowledge (as discussed in section 7.2.3), and the manner in which we designed the measuring instrument with respect to methodology. Since this was an original study, it was impossible to foresee, and therefore control a number of unpredictable variables. Our expectation was that modification to the methodology would be required for future work of this kind. Possible modifications are suggested below in relation to a

number of aspects.

7.5.2.1 Test Items

First, the number of test items was very small in the current study and thus reduced the reliability of the results and hence the validity of the research. Therefore, a vital modification in a repeated study would be the elicitation of more data by means of more test items. This concern was especially important for the correlational analysis, which requires at least 30 items to reach a secure level of reliability. However, owing to the limited reading speed of the informants (as mentioned in section 6.2.4), this was impossible in the current study. It is suggested that a higher level of informants (such as English major students), or sampling from a English course focusing on reading, may be able to satisfy this data sufficiency requirement.

7.5.2.2 Informants

Because informants of higher proficiency level were considered more appropriate for research designed to test the ability to recognize the structure of semantic categories, this research set limits on the proficiency level of the informants. For future research, a group of mixed levels might reveal more interesting findings with respect to the critical level at which L2 learners are able to recognize the structure of semantic categories. Moreover, sampling from native speakers can also be considered. Their categorization behaviours and reading mechanisms can be compared with those of non-native speakers so as to establish a possible norm.

7.5.2.3 Test Materials

It would be useful to know whether the type of test materials might affect the results of the experiments conducted in this study. In the present research, we used

graded readers as materials because they were currently used in the experimental setting and because they have recently affected considerable pedagogic interest in second language learning. For further research, it may be interesting to draw test items from different types of materials, such as newspapers or journals. In addition, the genres of graded readers and writing styles of different authors may be significant variables which were not investigated in this research, since each in liable to employ different ways of constructing semantic categories with consequential implications for the efficiency with which readers can re-construct these categories.

7.5.2.4 Task Design

In eliciting judgements as to inclusion properties of test items (Task 1), we did not include non-members in the examples presented to informants. If non-members were included, this would show whether the judgement of informants was influenced by reading experience, world knowledge or even personal experience. Needless to say, a more complicated scoring system would be required to code the results. In addition, informants might need more training in recognition of a category member than the short training section provided in the current study. In Task 2, in which informants were required to bracket category members, they sometimes bracketed only the verb of the target item, rather than the complete phrase. As a result, it was not clear whether they could recognize category members and therefore these answers were taken as incorrect. Gentner & Namy's study (1999) suggested a way to improve the design by giving an example in advance so as to demonstrate how to do the task, especially for categories containing more than one members.

241

7.5.2.5 Rhetorical Coupling

The perception of rhetorical coupling by non-native speakers appears to be a problematic issue that was not dealt with properly in this study. As the results showed, the non-native speakers seemed not to be able to perceive the rhetorical properties of the test items and therefore their performance did not follow our assumption that more diverse rhetorical properties entailed a stronger ability to discriminate the degree of category membership. One way to modify the design is to ask non-native speaker informants to rate these items first, and then set the task of differentiation of degree of category membership. As a consequence, we would be able to see more explicitly whether there is a close association between awareness of rhetorical properties and the processability of linguistic message. Such a method might be expected to address more directly the issue that a more salient message form will be liable to obscure message content. The present study only obtained rating data from native speakers and assumed that non-native speakers, especially learners of higher proficiency, could perceive the rhetorical properties in more or less the same way as the native speakers. But, since rhetorical properties are concerned with suprasegmental features, it may well be that there are individual differences between learners as well as a degree of difference between native speakers and learners.

7.5.2.6 Setting

Finally, the research may be replicated in a very different setting, for example, a more-controlled, laboratory-like course. In this research, we admitted that there were inherent unpredictable variables in the classroom setting. For example, since the reading of graded readers was only a small part of the course requirement, we were not certain about whether or not other requirements, such as the components of

coursebook study and English conversation practice, would affect the performance of these informants. Therefore, a setting in which reading played a moare central role might yield different results.

Despite these admitted limitation, the methodological contribution of this research deserves mentioning. Especially, the tasks that were designed to measure vocabulary depth knowledge were diverse and closely related to the reading experience of the informants. Therefore, these task types might be worth adopting in foreign/second language reading instruction. In addition, since the ability to recognize contextually determined semantic categories was linked directly to successful reading comprehension, the tasks we designed to measure vocabulary depth knowledge might thus be useful as reading comprehension check.

Appendix A: Personal Information

Please write down your personal information:

Name:

ID No:

Sex:

Major:

Age:

Education:

Years of learning English:

Appendix B: Consent Form

Dear Students:

You are involved in an experiment concerning vocabulary learning. The reader *"Trumpet Voluntary"* is specially chosen for this purpose.   The tests you have taken can provide valuable data for this experiment.   I appreciate your cooperation and need your consent to use the data to do further analysis.   Any result will be useful for vocabulary learning and teaching.   Please sign your names (both English and Chinese) below to give me permission to use the data.

With regards,

_____

Fei-Hsuan Liao

I agree that the researcher can use the data for his/her experiment.

_____

Appendix C: Categorization in Chinese

The nature of linguistic categorization in the mother tongue is worth mentioning as it may play a role in the informants' ability to recognize the inclusion properties of semantic categories. One overt means of linguistic categorization is the classifier systems found in Mandarin and other south-east Asian languages. Classifier systems typically sort members of a category on the basis of the different properties of objects, such as their shape (their one or two-dimensional extension) or their functional interaction with other objects. For instance, Mandarin employs "*tiao*" to classify long objects such as cucumbers, as in "*yi tiao huanggua*", towels as in "*yi tiao maojin*", and even boats as in "*yi tiao chuan*". As an illustration of a more abstract notion, movable objects, especially those with handles, are classified as a group with the classifier "*ba*" (literally "handle") as in "*yi ba yizi*" (a chair) and in "*yi ba jenda*" (a pair of scissors). The seemingly arbitrary use of classifiers is in effect highly cognitively-based, as the analysis of classifier system in Dyirbal, which groups women, fire and dangerous things as one class, has been shown to be (Lakoff, 1986). However, in English, classifier use is both rare and comparatively arbitrary. Thus, although *a piece of advice, a piece of cake*, and *a piece of furniture* may appear to designate a class, in fact "advice", "cake", and "furniture" are not perceived to have anything in common.

Although the cognitive base of the classifier system in Mandarin has salience for users of the language, the use of classifiers in Mandarin may not follow the prototypical view of categorization. At first sight, taking a particular classifier as a superordinate notion does not readily lead one to think of prototypical category members of the kind identified in Rosch's research (1973a, 1973b, 1975). This is because the categorization system reflects the diachronic development of Mandarin

through stages from particular lexical usage to generalization, so that over time different perceptual properties become the basis for categorization. This allows new members to be grouped in a class although they may have little in common with original members (Erbaugh, 1986: 426-28). Thus, a classifier such as *tiao* or *ba* used to group only prototypical members, but the expansion in the number of members has gradually resulted in the use of the classifier becoming conventionalized.

For example, the very common classifier *"zhang"* was used only for bows in Chinese in the Han Dynasty around 200 B.C. The Chinese character still reflects the object by combining the characters for "bow" and "long" (Erbaugh, 1986: 429). However, in due course, "zithers" were included as clear members because they were made of bows with stretchable strings. Thus, the notions "stretch" and "expand" were prototypical for the class at this time (Tai & Chao, 1994: 70-71). Historical developments have involved several stages of intensional expansion, so that *"zhang"* now covers stretchable and flat items such as paper, paintings, tables and even chairs, and the notion of "flat" has therefore become the prototypical notion for the classifier *"zhang"* in modern Mandarin, although the meaning of "flat" is difficult to reconcile with the original meaning of "bowed" and "curved". The category members to which this classifier applies are thus so divergent that it is no longer possible to identify prototypical members.

As a consequence, not only is one classifier used to cover a large number of divergent members, but also various classifiers may be found with a single noun. Thus, chair occurs with *"ba"*, as explained above, and also with *"zhang"*. That is, variable classifier choice is allowed. This occurs especially when a new invention brings a new item into the lexicon. Several classifiers emphasizing different

perceptual or functional properties may compete to include the new word. For example, at present the word "computer" occurs with a classifier such as *"bu"* as in *"yi bu zidian"* (a volume of dictionary), and with *"tai"* as in *"yi tai gichi"* (a machine).

These examples of noun classifiers illustrate that the Chinese classifiers include both perceptually similar and substantially different objects within the same categories. Although the basis for inclusion may seem to be relatively arbitrary for non-Chinese speakers, despite the forgoing discussion, for Chinese speakers most members of a category do possess perceptual similarity. In addition, the diachronic development of the classifiers partly accounts for the apparent inclusion of more divergent members into a category. In other words, we can say that the conceptual categorization of objects in Mandarin is very different from the kind of conceptual classification found in Western languages since Chinese makes overt morphological reference to the superordinate notion which determines category membership. For these reasons, the ability to recognize the inclusion properties of categories as revealed by this group of Chinese informants may (or may not) be facilitated or impeded by the conventional classifier system of their mother tongue and their particular experience of the world as reflected in it. Therefore, a comparison of the categorization behaviours of classifier and non-classifier language speakers in relation to semantic categories of the instantial kind investigated in this research is likely to be revealing.

Appendix D: Sample Text (Chapter 3, original text)[1]

Tibor and Malgosia became lovers, of course. Even when I think about that now it hurts. After all this time it still hurts. But there was nothing I could have done, even if I had known how to. She fell for him almost immediately. Perhaps it was the scene in the canteen that did it for her. Perhaps it was later in some pub or cafe that she looked at him and felt the joyous pain that love can be. I don't know. I only know that when I realised what the situation was I thought I would die with unhappiness. Because I was crazy about her too, absolutely obsessed with the thought of her. And though we were the best of friends, and she came to cry on my shoulder every time Tibor ignored her or cheated on her, and even though I was sure she was very fond of me, still she didn't love me like I, poor fool, loved her. Instead I had to watch her eyes light up with excitement every time Tibor walked into the room. I had to learn to grin and bear it every time she cancelled one of our visits to the cinema or a folk club because she was going to be with him. I had to accept that, compared to Tibor, I would always be second best.

One afternoon, a few weeks into that first term, I went into Duke's Hall because I could hear the sound of a trumpet. I pushed open the glass doors. Malgosia was on the empty stage making the most beautiful noise I had ever heard. Another student was accompanying her on the organ. They were playing the 'Trumpet Voluntary', one of the most famous English trumpet tunes there is.

I stood and listened, until she had finished. Then she saw me, and smiled.

'That was beautiful,' I told her. 'I love Purcell's music.'

'It's not Purcell,' Malgosia said. 'Everybody thinks he wrote it, but it was actually written by someone called Jeremiah Clarke. I've been reading about it. Did you know,' she went on, 'he killed himself for love. It's the kind of thing they did in those days. Isn't it romantic?'

'No,' I laughed. 'It's stupid if you ask me.'

'Ah,' she replied dreamily, 'you are just a man, an English man. Did you know,' she went on, 'he killed himself for love. It's the kind of thing they did in those days. Isn't it romantic?'

'No,' I laughed. 'It's stupid if you ask me.'

'Ah,' she replied dreamily, 'you are just a man, an English man. But to die for love! That takes a more passionate soul than yours perhaps.'

*   *   *

At the end of the first year, when the summer holidays began, Malgosia went back to Poland to be with her family. Right up until the moment she left I had this dream that she would invite me to visit her there, but somehow I knew that she wouldn't. My parents, who knew nothing of my romantic depression, wanted me to go on holiday with them, but with the arrogance of youth I couldn't stand the idea. I was much happier busking on the streets of London, playing for cinema queues and shopping tourists with the new student quartet I had joined at the end of the summer term. Our first violin player, Carl Robinson, was rather uneasy about playing Mozart's 'Eine Kleine Nachtmusik' ten times a day (because he'd rather be attempting Shostakovich and Bartok) and Rachel Merino, our cellist, was uncomfortable with the long metal spike which allowed her to play standing up.

---

[1] The reason for presenting the whole text, rather than only test items, is to provide a more complete idea of the wider context so as to enable the reader to consider whether the most appropriate items were selected for the experiment.

'Still,' as Matt Jenkins, the second violinist said, 'it's good performance practice, and we're earning quite a bit of money too. No-one's complaining about that!'

It was true. People seemed to like the way we played, and my open viola case soon filled up with small change. We earned even more when a law student who lived with me came to help us by going round our audiences with a hat, asking for donations. Ken, a Nigerian, wasn't very keen on the music we played but he was a magician with the hat. He almost ran around the crowds that were listening to us, smiling at them, laughing and joking, telling them we were poor refugees, anything to make them put their hands in their pockets. It looked like it was going to be a good summer. Sometimes I even forgot about Malgosia. But not for long.

One evening, as we packed up our instruments for the day, Rachel suggested that we go to a pub by the river in Chiswick where she lived. The next day was Sunday so we weren't going to play, and anyway we needed a break. The others didn't want to go all that way because it's a long way from Central London but I had nothing better to do, so I agreed.

It was a lovely evening so Rachel and I sat outside and watched the life of the river in front of us. There were birds, the sounds of a great city all around us, planes making their way westwards to Heathrow airport, rowers shouting orders at each other, and a little police boat making its way down river, its blue light flashing as it went as fast as it could towards some emergency.

I liked Rachel. She was quiet and gentle. She had light brown hair, and pretty, brown eyes set in a round, pleasant face. When she smiled she looked like a happy child and you knew you could trust her. She was very easy to talk to.

That night we sat and talked about what we hoped for the future. I told her I wanted to make enough money as a musician to have a nice house, travel a bit, that kind of thing. She told me that her dreams were much the same. She wanted children one day, she said, but for that she'd need to find the right man.

'Well it's no good looking at me,' I said, as a joke.

'I know that, you fool,' she said, laughing at me. 'You can't see anybody anyway. Not while Malgosia is in the way.' I blushed.

'Maybe,' I replied. I didn't like talking about it.

'Can I say something?' Rachel asked, nervously.

'It depends what it is,' I replied. Around us people were talking and laughing as the night got darker. I saw the lights of a party boat travelling along the river in front of us.

'It's just that, well, I know Malgosia is beautiful. I mean really beautiful. I wish I was beautiful like that. And I do like her. But she's crazy about Tibor, and anyone who's crazy about Tibor, well É' She stopped and looked at me, wondering how I would react.

'Well what?' I answered. I understood what she was saying, I think, but I didn't like anyone criticising Malgosia.

'Oh God, now you're cross with me,' Rachel worried. 'Sorry. Sorry. But it's just a pity to see you and her. She's not right for you. You're wasting your time, wasting your life on her and you're not getting anything back. It doesn't look good. That's what I think.'

'Well,' I snapped back, without thinking, 'I don't care what you think, OK? Me and Malgosia, well, we're É' I wanted a word that meant more than 'friend' but I couldn't think of the right one – 'we're special, all right? So it's none of your business. Just keep out of my affairs, OK?'

Rachel had gone red and I had gone too far. My only excuse is that I was very

confused then, and still very young. But I suppose, if I am honest, that wasn't it. It was because Rachel had said something that I didn't want to hear because it was the truth. Now I think that if only I had listened to her then, if only I had understood what she was trying to tell me, I might not have made the decisions that I did and my life might have turned out very differently.

We finished our drinks in silence. I tried to start conversations again once or twice, but she only answered with 'yes' or 'no', so it wasn't much good. In the end, when closing time was called by the pub landlord, we just said goodbye and I caught the tube back into London.

As the train made its way noisily through the darkness I thought about the conversation I had just had and I realised that it probably meant the end of the quartet.

'Oh well,' I thought, 'I can join another one, or we'll find a new cellist. It doesn't matter.'

And then I suddenly saw Rachel's face, Rachel's sad gentle face, and perhaps it was the drink, or perhaps it was because it was a warm summer's night, or perhaps I was just lonely, I don't know, but anyway, at the next station I got off the train, walked over the bridge and got on the first train which was going back the way I had just come.

Half an hour later I was at the door of Rachel's house. I rang the bell. The door opened.

'Derek,' Rachel said in surprise, 'I thought you had gone home.'

'I was going to,' I replied, 'and then I just thought É' but I didn't really know what I had thought so I kissed her instead. She didn't seem to mind so I kissed her again, and it was very nice and soon just kissing didn't seem to be enough and, well, you can imagine the rest.

The next morning we couldn't look at each other. We had gone too far, too fast. All I could think was that I had somehow betrayed Malgosia (which was ridiculous, since she didn't love me), and all Rachel could think was that I had used her.

It was raining when I reached the tube station. I didn't feel good and wondered how to cheer myself up. When I got back to my house the first thing Ken said to me was, 'Where have you been? You look awful.'

'Thanks,' I said, 'thanks a lot.'

'Hey, don't you worry about it. You can sleep it off, whatever it was' Ken said.

'I don't know,' I answered.

'Well, how about a coffee while you're thinking about it?' he offered, so we sat and talked and I told him about Rachel, even though I knew I shouldn't. But I needed someone to talk to.

'Well,' he said when I had finished my story of the previous night, 'things like that don't happen to me over here. Not yet anyway. You're just lucky.'

'Then why don't I feel lucky?' I asked him.

'Well now,' he laughed, 'I can't help you there. Maybe you're crazy or something. Yes, that's probably it,' and we laughed, and after half an hour in his company I felt better because he didn't seem to think that I had done anything terrible, and because he let me tell him about Malgosia and he didn't say that I was wasting my time.

The next day the quartet met up outside Covent Garden tube station as we had done every day since we started busking. I wasn't sure whether Rachel would be there, but she had already arrived when I got out of the lift which had brought me up from the depths of London and walked out into the sunshine. She avoided my eyes and

wouldn't talk to me at first, but later, when we were setting up at one end of the old market building and the other two were getting their violins out of their cases, she came up to me.

'Derek,' she said, 'can we have a quick word?'

'Yes,' I said nervously.

'It's just this,' she announced seriously, looking away from me, 'what happened happened, I know that, but it's not going to happen again, is it?'

'No,' I said, and I meant it.

'Right,' she said in a controlled way. 'And you want the quartet to continue, don't you?'

'Yes. Yes, I think so.'

'So can we just be friends?' she asked. I wonder how much that cost her.

'If that's OK,' I said. I felt a great sense of relief, to be honest. I had expected more problems than this.

'Yes, it is OK,' she said and started to walk away. But then she turned back, came right up to me, and for the first time that day she looked me full in the eyes.

'One last thing,' she whispered. 'I don't want anyone else knowing about our É knowing about it. And we won't talk about it ever again. OK?'

'OK,' I agreed. She relaxed then, smiled at me, walked away again and before long we were halfway through the first movement of the Mozart and the money was falling into Ken's hat.

But one afternoon two weeks later when we broke for lunch, Rachel refused to go on playing, and she wouldn't say why, and that was the end of our little band. At least for the time being.

We watched her walking away, her head down. She hadn't said goodbye to anyone.

'What was that all about?' Carl asked as Rachel disappeared round a corner. 'Does anyone know what's going on?'

We all looked at each other. Nobody had any idea.

'What about you, Ken?' Matt said, as our money-collector came towards us with four pints of beer on an old bar tray from the pub behind us.

'What about what?' Ken answered.

'Why did Rachel go off like that? Do you know?'

'Me?' said Ken. 'Why should I know? Don't ask me. Must be some musician thing. Why are you asking me?'

'Hey,' I told him, 'don't be so defensive. I'm sure it wasn't your fault.' But of course it was. His fault and my fault. Rachel and Ken had been walking towards the pub together (Ken told me later), and he'd asked her how she felt about me now. And Rachel had said, 'What do you mean?' and Ken had answered, 'You know, after that night of yours, that night together?' And Rachel had stopped in the middle of the street, gone bright red, and said, 'Who told you that? How do you know that? Did Derek tell you that?' her voice getting louder with every question. Then she'd marched past Ken, packed up her cello, told us she was leaving, and walked away without a word of farewell.

<p style="text-align:center">* * *</p>

The summer holidays ended and the autumn term at the Academy began. There was no sign of Malgosia the first day, and she hadn't answered either of the letters I had sent her, so I was pretty miserable. I tried to talk to Rachel, but she refused to speak to me. I asked her, begged her to play with us again, and all she said was, 'How can I play with you again? You told Ken all about us, didn't you? Why should I spend any more time in your company than I have to?'

Things weren't going well with my studies either. My viola teacher was ill and the man who took her place for that first week didn't like me and I felt the same about him. My piano playing wasn't getting much better either. Carl and Matt wanted to get another cellist for our quartet, but somehow I could not agree to that. I had already done enough damage there, and I still hoped, one day, that Rachel would change her mind. And Ken wouldn't talk to me because I had lost my temper with him about what he had said to Rachel.

So I suppose you could say that my life was a mess. I was playing badly. I didn't have many friends, and the person I thought I loved had disappeared from my life. But not for long. One evening, six days after term started, I was sitting in my room watching a bad old Hollywood movie on my small television when someone knocked at my door. I opened it. Ken stood outside.

'Hi,' I said, 'come in. You haven't been around for a few days. Where have you been?'

'Don't worry about where I've been,' he replied, not looking me in the eye. 'You've got more important things to think about.'

'What do you mean?'

'There's someone downstairs to see you,' he said, 'and she doesn't look very happy.'

'Who is it?' I asked him.

'I don't know.' he replied, 'I've never seen her before, but she's really lovely. If you're not interested, just let me know. I'd love to try and cheer her up.'

I left him and walked downstairs. The front door was still open. It was beginning to get dark outside and a heavy rain was falling. Malgosia was standing in the passageway with two suitcases, her beautiful hair dripping wet and a look of complete misery on her face.

Appendix E: Sample Text (Chapter 3, modified text)

Tibor and Malgosia became lovers, of course. Even when I think about that now it hurts. After all this time it still hurts. But there was nothing I could have done, even if I had known how to. She fell for him almost immediately. Perhaps it was the scene in the canteen that did it for her. Perhaps it was later in some pub or cafe that she looked at him and felt the joyous pain that love can be. I don't know. I only know that when I realised what the situation was I thought I would die with unhappiness. Because I was crazy about her too, absolutely obsessed with the thought of her. And though we were the best of friends, and she came to cry on my shoulder every time Tibor ignored her or cheated on her, and even though I was sure she was very fond of me, still she didn't love me like I, poor fool, loved her. Instead I had to watch her eyes light up with excitement every time Tibor walked into the room. I had to learn to grin and bear it every time she cancelled one of our visits to the cinema or a folk club because she was going to be with him. I had to accept that, compared to Tibor, I would always be second best.

One afternoon, a few weeks into that first term, I went into Duke's Hall because I could hear the sound of a trumpet. I pushed open the glass doors. Malgosia was on the empty stage making the most beautiful noise I had ever heard. Another student was accompanying her on the organ. They were playing the 'Trumpet Voluntary', one of the most famous English trumpet tunes there is.

I stood and listened, until she had finished. Then she saw me, and smiled.

'That was beautiful,' I told her. 'I love Purcell's music.'

'It's not Purcell,' Malgosia said. 'Everybody thinks he wrote it, but it was actually written by someone called Jeremiah Clarke. I've been reading about it. Did you know,' she went on, 'he killed himself for love. It's the kind of thing they did in those days. Isn't it romantic?'

'No,' I laughed. 'It's stupid if you ask me.'

'Ah,' she replied dreamily, 'you are just a man, an English man. Did you know,' she went on, 'he killed himself for love. It's the kind of thing they did in those days. Isn't it romantic?'

'No,' I laughed. 'It's stupid if you ask me.'

'Ah,' she replied dreamily, 'you are just a man, an English man. But to die for love! That takes a more passionate soul than yours perhaps.'

*   *   *

At the end of the first year, when the summer holidays began, Malgosia went back to Poland to be with her family. Right up until the moment she left I had this dream that she would invite me to visit her there, but somehow I knew that she wouldn't. My parents, who knew nothing of my romantic depression, wanted me to go on holiday with them, but with the arrogance of youth I couldn't stand the idea. I was much happier busking on the streets of London, playing for cinema queues and shopping tourists with the new student quartet I had joined at the end of the summer term. Our first violin player, Carl Robinson, was rather uneasy about playing Mozart's 'Eine Kleine Nachtmusik' ten times a day (because he'd rather be attempting Shostakovich and Bartok) and Rachel Merino, our cellist, was uncomfortable with the long metal spike which allowed her to play standing up.

'Still,' as Matt Jenkins, the second violinist said, 'it's good performance practice, and we're earning quite a bit of money too. No-one's complaining about that!'

It was true. People seemed to like the way we played, and my open viola case soon filled up with small change. We earned even more when a law student who lived

with me came to help us by going round our audiences with a hat, asking for donations. Ken, a Nigerian, wasn't very keen on the music we played *but he was a magician with the hat in the way that he almost ran around the crowds that were listening to us, smiling at them, laughing and joking, telling them we were poor refugees, anything to make them put their hands in their pockets.* It looked like it was going to be a good summer. Sometimes I even forgot about Malgosia. But not for long.

One evening, as we packed up our instruments for the day, Rachel suggested that we go to a pub by the river in Chiswick where she lived. The next day was Sunday so we weren't going to play, and anyway we needed a break. The others didn't want to go all that way because it's a long way from Central London but I had nothing better to do, so I agreed.

It was a lovely evening so *Rachel and I sat outside and watched the life of the river in front of us with its birds, the sounds of a great city all around us, planes making their way westwards to Heathrow airport, rowers shouting orders at each other, and a little police boat making its way down river, its blue light flashing as it went as fast as it could towards some emergency.*

I liked Rachel because she was quiet and gentle and had light brown hair, and pretty, brown eyes set in a round, pleasant face and when she smiled she looked like a happy child and you knew you could trust her; and she was very easy to talk to. That night we sat and talked about what we hoped for the future so I told her I wanted to make enough money as a musician to have a nice house, travel a bit, that kind of thing. She told me that her dreams were much the same. She wanted children one day, she said, but for that she'd need to find the right man.

'Well it's no good looking at me,' I said, as a joke.

'I know that, you fool,' she said, laughing at me. 'You can't see anybody anyway. Not while Malgosia is in the way.' I blushed.

'Maybe,' I replied. I didn't like talking about it.

'Can I say something?' Rachel asked, nervously.

'It depends what it is,' I replied. Around us people were talking and laughing as the night got darker. I saw the lights of a party boat travelling along the river in front of us.

'It's just that, well, I know Malgosia is beautiful. I mean really beautiful. I wish I was beautiful like that. And I do like her. But she's crazy about Tibor, and anyone who's crazy about Tibor, well É' She stopped and looked at me, wondering how I would react.

'Well what?' I answered. I understood what she was saying, I think, but I didn't like anyone criticising Malgosia.

'Oh God, now you're cross with me,' Rachel worried. 'Sorry. Sorry. But it's just a pity to see you and her. She's not right for you. You're wasting your time, wasting your life on her and you're not getting anything back. It doesn't look good. That's what I think.'

'Well,' I snapped back, without thinking, 'I don't care what you think, OK? Me and Malgosia, well, we're É' I wanted a word that meant more than 'friend' but I couldn't think of the right one – 'we're special, all right? So it's none of your business. Just keep out of my affairs, OK?'

Rachel had gone red and I had gone too far. My only excuse is that I was very confused then, and still very young. But I suppose, if I am honest, that wasn't it. It was because Rachel had said something that I didn't want to hear because it was the truth. Now I think that if only I had listened to her then, if only I had understood

what she was trying to tell me, I might not have made the decisions that I did and my life might have turned out very differently.

We finished our drinks in silence. I tried to start conversations again once or twice, but she only answered with 'yes' or 'no', so it wasn't much good. In the end, when closing time was called by the pub landlord, we just said goodbye and I caught the tube back into London.

As the train made its way noisily through the darkness I thought about the conversation I had just had and I realised that it probably meant the end of the quartet.

'Oh well,' I thought, 'I can join another one, or we'll find a new cellist. It doesn't matter.'

And then I suddenly saw Rachel's face, Rachel's sad gentle face, and perhaps it was the drink, or perhaps it was because it was a warm summer's night, or perhaps I was just lonely, I don't know, but anyway, at the next station I got off the train, walked over the bridge and got on the first train which was going back the way I had just come.

Half an hour later I was at the door of Rachel's house. I rang the bell. The door opened.

'Derek,' Rachel said in surprise, 'I thought you had gone home.'

'I was going to,' I replied, 'and then I just thought É' but I didn't really know what I had thought so I kissed her instead. She didn't seem to mind so I kissed her again, and it was very nice and soon just kissing didn't seem to be enough and, well, you can imagine the rest.

The next morning we couldn't look at each other. We had gone too far, too fast. All I could think was that I had somehow betrayed Malgosia (which was ridiculous, since she didn't love me), and all Rachel could think was that I had used her.

It was raining when I reached the tube station. I didn't feel good and wondered how to cheer myself up. When I got back to my house the first thing Ken said to me was, 'Where have you been? You look awful.'

'Thanks,' I said, 'thanks a lot.'

'Hey, don't you worry about it. You can sleep it off, whatever it was' Ken said.

'I don't know,' I answered.

'Well, how about a coffee while you're thinking about it?' he offered, so we sat and talked and I told him about Rachel, even though I knew I shouldn't. But I needed someone to talk to.

'Well,' he said when I had finished my story of the previous night, 'things like that don't happen to me over here. Not yet anyway. You're just lucky.'

'Then why don't I feel lucky?' I asked him.

'Well now,' he laughed, 'I can't help you there. Maybe you're crazy or something. Yes, that's probably it,' and we laughed, and after half an hour in his company I felt better because he didn't seem to think that I had done anything terrible, and because he let me tell him about Malgosia and he didn't say that I was wasting my time.

The next day the quartet met up outside Covent Garden tube station as we had done every day since we started busking. I wasn't sure whether Rachel would be there, but she had already arrived when I got out of the lift which had brought me up from the depths of London and walked out into the sunshine. She avoided my eyes and wouldn't talk to me at first, but later, when we were setting up at one end of the old market building and the other two were getting their violins out of their cases, she came up to me.

'Derek,' she said, 'can we have a quick word?'

'Yes,' I said nervously.

'It's just this,' she announced seriously, looking away from me, 'what happened happened, I know that, but it's not going to happen again, is it?'

'No,' I said, and I meant it.

'Right,' she said in a controlled way. 'And you want the quartet to continue, don't you?'

'Yes. Yes, I think so.'

'So can we just be friends?' she asked. I wonder how much that cost her.

'If that's OK,' I said. I felt a great sense of relief, to be honest. I had expected more problems than this.

'Yes, it is OK,' she said and started to walk away. But then she turned back, came right up to me, and for the first time that day she looked me full in the eyes.

'One last thing,' she whispered. 'I don't want anyone else knowing about our É knowing about it. And we won't talk about it ever again. OK?'

'OK,' I agreed. She relaxed then, smiled at me, walked away again and before long we were halfway through the first movement of the Mozart and the money was falling into Ken's hat.

But one afternoon two weeks later when we broke for lunch, Rachel refused to go on playing, and she wouldn't say why, and that was the end of our little band. At least for the time being.

We watched her walking away, her head down. She hadn't said goodbye to anyone.

'What was that all about?' Carl asked as Rachel disappeared round a corner. 'Does anyone know what's going on?'

We all looked at each other. Nobody had any idea.

'What about you, Ken?' Matt said, as our money-collector came towards us with four pints of beer on an old bar tray from the pub behind us.

'What about what?' Ken answered.

'Why did Rachel go off like that? Do you know?'

'Me?' said Ken. 'Why should I know? Don't ask me. Must be some musician thing. Why are you asking me?'

'Hey,' I told him, 'don't be so defensive. I'm sure it wasn't your fault.' But of course it was. His fault and my fault. Rachel and Ken had been walking towards the pub together (Ken told me later), and he'd asked her how she felt about me now. And Rachel had said, 'What do you mean?' and Ken had answered, 'You know, after that night of yours, that night together?' And Rachel had stopped in the middle of the street, gone bright red, and said, 'Who told you that? How do you know that? Did Derek tell you that?' her voice getting louder with every question. Then she'd marched past Ken, packed up her cello, told us she was leaving, and walked away without a word of farewell.

* * *

The summer holidays ended and the autumn term at the Academy began. There was no sign of Malgosia the first day, and she hadn't answered either of the letters I had sent her, so I was pretty miserable. I tried to talk to Rachel, but she refused to speak to me. I asked her, begged her to play with us again, and all she said was, 'How can I play with you again? You told Ken all about us, didn't you? Why should I spend any more time in your company than I have to?'

Things weren't going well with my studies either since my viola teacher was ill and the man who took her place for that first week didn't like me and I felt the same about him and my piano playing wasn't getting much better either. Carl and Matt

wanted to get another cellist for our quartet, but somehow I could not agree to that. I had already done enough damage there, and I still hoped, one day, that Rachel would change her mind. And Ken wouldn't talk to me because I had lost my temper with him about what he had said to Rachel.

So I suppose you could say that my life was a mess since I was playing badly and didn't have many friends, and the person I thought I loved had disappeared from my life.

But not for long. One evening, six days after term started, I was sitting in my room watching a bad old Hollywood movie on my small television when someone knocked at my door. I opened it. Ken stood outside.

'Hi,' I said, 'come in. You haven't been around for a few days. Where have you been?'

'Don't worry about where I've been,' he replied, not looking me in the eye. 'You've got more important things to think about.'

'What do you mean?'

'There's someone downstairs to see you,' he said, 'and she doesn't look very happy.'

'Who is it?' I asked him.

'I don't know.' he replied, 'I've never seen her before, but she's really lovely. If you're not interested, just let me know. I'd love to try and cheer her up.'

I left him and walked downstairs. The front door was still open. It was beginning to get dark outside and a heavy rain was falling. Malgosia was standing in the passageway with two suitcases, her beautiful hair dripping wet and a look of complete misery on her face.

9 November 2001

Dear Friend,

This is to ask for your help in my research into vocabulary learning by providing your judgement on each of the extracts on the attached sheet. These extracts are taken from a reader for advanced language learners. Amongst other things, I am trying to find out whether the degree of rhetorical patterning present in a text affects vocabulary learning in any way. In order to do this, I need the help of native speakers such as yourself to decide on the extent to which rhetorical patterning exists in a number of short extracts which I am planning to use as test items.

On the back of this letter you will find a training exercise. After completing this exercise now, all you need to do later is to circle the score you give each extract on the attached sheet and bring the sheet back to class next week.

Thank you in advance for your help, which is very much appreciated.

Frances Liao

Study the italicized part of each of the following extracts which have been taken from a reader for advanced learners. Each italicized fragment contains two concepts. Please decide on the degree of rhetorical patterning which they share on a four point scale:

O = there is no obvious rhetorical parallel
1 = there is some rhetorical coupling
2 = there is a distinct level of rhetorical coupling
3 = the structures are (virtually) parallel

and indicate your rating by circling the appropriate number on the line below the extract.

(For example, most people would agree that *M 'n' S* exhibits a greater degree of rhetorical coupling than *Marks and Spencer* but not as great a degree as *Marks and Sparks*.)

Example 1

..that would land us both in *a lot of trouble, a great deal of trouble.*

/3/ /2/ /1/ /0/

Example 2

There was a great bang and something hit me hard on the side of my head. *I felt my legs go weak and the day went all dark on me.*

/3/ /2/ /1/ /0/

Example 3

The comfortable dark was asking for me again. *The light faded, the voice disappeared.*

/3/ /2/ /1/ /0/

Example 4

"*Your Malgosia isn't here anymore*, that's what," Oswaldo said bluntly. "*She's gone.*"

/3/ /2/ /1/ /0/

Study the italicized part of each of the following extracts which have been taken from a reader for advanced learners. Each italicized fragment contains at least two concepts. Please decide on the degree of rhetorical patterning which they share on a four point scale:

O = there is no obvious rhetorical parallel
1 = there is some rhetorical coupling
2 = there is a distinct level of rhetorical coupling
3 = the structures are (virtually) parallel

and indicate your rating by circling the appropriate number on the line below the extract.

1   they'd taken me to Krakow and Gdansk, driving over the long roads of this proud country with *its unbeaten spirit, keeping its identity*..

/3/ /2/ /1/ /0/

2   It made me realise where Malgosia got *her bright-eyed intelligence* and *her deep dark passion.*

/3/ /2/ /1/ /0/

3   Now I was in a taxi, unhappy and confused, *my head still bandaged, my mind trying to come to terms with all that had happened to me over the last week, my body tired and hurting, the terrible plane smell in my hair and on my clothes.*

/3/ /2/ /1/ /0/

4   Her hair was cut short. She wore heavy black glasses. *She watched me as I struggled in with my suitcase. She didn't look welcoming at all.*

/3/ /2/ /1/ /0/

5   'I feel terrible, all right?' I replied angrily. *'My wife's run off with someone else, I've been questioned by the police, I've travelled halfway round the world, I've been shot at and nearly killed. Now my wife's supposed to be very ill and I don't know why, or where she is,*

but I know it's pretty damn serious.

/3/ /2/ /1/ /0/

6   'Why not?' I shouted at her, 'why not?'
'Because I don't dare. Because I can't bear it. Because she was my best friend before Tibor.'

/3/ /2/ /1/ /0/

7   every time he came here ... for years ... and then *she took him away from me. She took him back.*

/3/ /2/ /1/ /0/

8   It took a bit of time *for his words to sink in.* It took me time *to realise that Malgosia couldn't hear me in there.* It took me time *to realise that after all my journeying I had finally found her and it wouldn't do me any good at all.*

/3/ /2/ /1/ /0/

9   I was *too tired to feel anything, too shocked to understand what had happened, too alone to feel any warmth from the people around me.*

/3/ /2/ /1/ /0/

10 Anja was *nowhere to be seen.* Nobody mentioned *her absence.*

/3/ /2/ /1/ /0/

11 there was a blank space in my brain *where nothing moved, nothing was happening.*

/3/ /2/ /1/ /0/

12   Malgosia's father managed to ask his questions, though. *'What kind of a husband do you call yourself?'* was his first effort. *'What kind of a husband that lets this happen?'*

/3/ /2/ /1/ /0/

13   'Poisoned?' I repeated.   *I hadn't the slightest idea what he was talking about.*
'You didn't know? You didn't know about this?'

/3/ /2/ /1/ /0/

14   Anja started up *a low terrible moaning which seemed to go on and on.* I wished she would just stop. Even when her father made me go with him into the kitchen we could still hear it, *the inhuman noise of someone who was going to feel guilty for ever.*

/3/ /2/ /1/ /0/

15   what Malgosia's father told me was that my wife had died, in the end, *bleeding unstoppably, completely unable to breathe,*

/3/ /2/ /1/ /0/

16   both typical symptoms of someone who had been poisoned by *some form of nerve gas, some kind of chemical agent.*

/3/ /2/ /1/ /0/

17   Though *they still have to do more tests. They haven't been able to identify what type it was yet.*

/3/ /2/ /1/ /0/

18 *All the doctor will say is* that he thinks she must have been exposed to some deadly chemical a few days ago. *That's his guess anyway.*'

/3/ /2/ /1/ /0/

19 I'm just trying *to absorb what you're telling me.* I'm trying *to come to terms with what's happened,*

/3/ /2/ /1/ /0/

20 but suddenly I could do nothing. *I felt my legs go from under me, and yet again, blackness washed over me like a wave and I passed out.*

/3/ /2/ /1/ /0/

21 The funeral was as bad as a funeral could be; *grey rain coming at us almost horizontally from a grey cold sky, the voice of the priest at the graveside taken away from us by a bitter wind, the tears of the mourners invisible in the wet cold.*

/3/ /2/ /1/ /0/

22 *It was a bumpy journey,* with turbulence almost all the way. Most of the time I half slept as *the plane bounced through the skies,*

/3/ /2/ /1/ /0/

23 When the plane landed *it taxied away towards the terminal building,* but then, unexpectedly, *it came to a stop.*

/3/ /2/ /1/ /0/

Thank you very, very much for your help. Please remember to bring this sheet back

to class on 16 November.

Appendix G: Average Rating for Each Item (high to low degree)

Item 16   both typical symptoms of someone who had been poisoned by *some form of nerve gas, some kind of chemical agent.* **(2.6)**

Item 9   I was *too tired to feel anything, too shocked to understand what had happened, too alone to feel any warmth from the people around me.* **(2.4)**

Item 6   'Why not?' I shouted at her, 'why not?'
'Because I don't dare. Because I can't bear it. Because she was my best friend before Tibor.' **(2.2)**

Item 7   every time he came here ... for years ... and then *she took him away from me. She took him back.* **(2.2)**

Item 12   Malgosia's father managed to ask his questions, though. *'What kind of a husband do you call yourself?'* was his first effort. *'What kind of a husband that lets this happen?'* **(2)**

Item 2   It made me realise where Malgosia got *her bright-eyed intelligence* and *her deep dark passion.* **(1.9)**

Item 3   Now I was in a taxi, unhappy and confused, *my head still bandaged, my mind trying to come to terms with all that had happened to me over the last week, my body tired and hurting, the terrible plane smell in my hair and on my clothes.* **(1.5)**

Item 5   'I feel terrible, all right?' I replied angrily. *'My wife's run off with someone else, I've been questioned by the police, I've travelled halfway round the world, I've been shot at and nearly killed. Now my wife's supposed to be very ill and I don't know why, or where she is, but I know it's pretty damn serious.* **(1.5)**

Item 19   I'm just trying *to absorb what you're telling me.* I'm trying *to come to terms with what's happened,* **(1.5)**

Item 21   The funeral was as bad as a funeral could be; *grey rain coming at us almost horizontally from a grey cold sky, the voice of the priest at the graveside taken away from us by a bitter wind, the tears of the mourners invisible in the wet cold.* **(1.4)**

Item 11   there was a blank space in my brain *where nothing moved, nothing was happening.* **(1.4)**

Item 8   It took a bit of time *for his words to sink in.* It took me time *to realise that Malgosia couldn't hear me in there.* It took me time *to realise that after all my journeying I had finally found her and it wouldn't do me any good at all.* **(1.1)**

Item 14   Anja started up *a low terrible moaning which seemed to go on and on.* I wished she would just stop. Even when her father made me go with him into the kitchen we could still hear it, *the inhuman noise of someone who was going to feel*

guilty for ever. (1.2)

Item 22   *It was a bumpy journey,* with turbulence almost all the way. Most of the time I half slept as *the plane bounced through the skies,* (0.9)

Item 1   they'd taken me to Krakow and Gdansk, driving over the long roads of this proud country with *its unbeaten spirit, keeping its identity..*(0.8)

Item 13   'Poisoned?' I repeated.   *I hadn't the slightest idea what he was talking about.
'You didn't know? You didn't know about this?'* (0.8)

Item 23   When the plane landed *it taxied away towards the terminal building,* but then, unexpectedly, *it came to a stop.* (0.8)

Item 17   Though *they still have to do more tests. They haven't been able to identify what type it was yet.* (0.7)

Item 10   Anja was *nowhere to be seen.* Nobody mentioned *her absence.* (0.6)

Item 20   but suddenly I could do nothing. *I felt my legs go from under me, and yet again, blackness washed over me like a wave and I passed out.* (0.6)

Item 15   what Malgosia's father told me was that my wife had died, in the end, *bleeding unstoppably, completely unable to breathe,* (0.5)

Item 18   *All the doctor will say is* that he thinks she must have been exposed to some deadly chemical a few days ago. *That's his guess anyway.'* (0.5)

Item 4   Her hair was cut short. She wore heavy black glasses. *She watched me as I struggled in with my suitcase. She didn't look welcoming at all.*(0.4)

Appendix H: Task 1, Part 1 (Experiment 1)

Part 1: Please answer the following questions.
Chapter 3

Question 1 *a messy life*

Tick either (a) or (b) below
a) I don't understand this phrase
b) I think I understand this phrase

If you have ticked (a), go straight to the next question
If you have ticked (b), you should also answer (c)

c) The kinds of thing I think of when I think of a *messy life* include
Example 1 ....
Example 2 ....

Question 2 *ways of being a magician with the hat*

Tick either (a) or (b) below
a) I don't understand this phrase
b) I think I understand this phrase

If you have ticked (a), go straight to the next question
If you have ticked (b), you should also answer (c)

c) The kinds of thing I think of when I think of *ways of being a magician with the hat* include
Example 1 ....
Example 2 ....

Question 3 *reasons for liking a friend*

Tick either (a) or (b) below
a) I don't understand this phrase
b) I think I understand this phrase
If you have ticked (a), go straight to the next question
If you have ticked (b), you should also answer (c)

c) The kinds of thing I think of when I think of *reasons for liking a friend* include
Example 1 ....
Example 2 ....

Question 4   *hopes for the future*

Tick either (a) or (b) below
a) I don't understand this phrase
b) I think I understand this phrase

If you have ticked (a), go straight to the next question
If you have ticked (b), you should also answer (c)

c) The kinds of thing I think of when I think of *hopes for the future* include
Example 1 ....
Example 2 ....

Question 5   *things that don't go well with one's studies*

Tick either (a) or (b) below
a) I don't understand this phrase
b) I think I understand this phrase

If you have ticked (a), go straight to the next question
If you have ticked (b), you should also answer (c)

c) The kinds of thing I think of when I think of *things that don't go well with one's studies* include
Example 1 ....
Example 2 ....

Question 6   *the life of the river*

Tick either (a) or (b) below

a) I don't understand this phrase
b) I think I understand this phrase

If you have ticked (a), go straight to the next question
If you have ticked (b), you should also answer (c)

c) The kinds of thing I think of when I think of *the life of the river* include
Example 1 ....
Example 2 ....

Chapter 4

Question 1   what happens later in a relationship when people are unhappy

Tick either (a) or (b) below
a) I don't understand this phrase
b) I think I understand this phrase

If you have ticked (a), go straight to the next question
If you have ticked (b), you should also answer (c)

c) The kinds of thing I think of when I think of *what happens later in a relationship when people are unhappy* include
Example 1 ....
Example 2 ....

Question 2   *having a dream*

Tick either (a) or (b) below
a) I don't understand this phrase
b) I think I understand this phrase

If you have ticked (a), go straight to the next question
If you have ticked (b), you should also answer (c)

c) The kinds of thing I think of when I think of *having a dream* include
Example 1 ....
Example 2 ....

Question 3   *all that had made Madonna the person she is*

Tick either (a) or (b) below
a) I don't understand this phrase
b) I think I understand this phrase

If you have ticked (a), go straight to the next question
If you have ticked (b), you should also answer (c)

c) The kinds of thing I think of when I think of *all that had made Madonna the person she is* include
Example 1 ....
Example 2 ....

Question 4   *a long conversation*

Tick either (a) or (b) below

a) I don't understand this phrase
b) I think I understand this phrase

If you have ticked (a), go straight to the next question
If you have ticked (b), you should also answer (c)

c) The kinds of thing I think of when I think of *a long conversation* include
Example 1 ....
Example 2 ....

Question 5 *having a comfortable life*

Tick either (a) or (b) below
a) I don't understand this phrase
b) I think I understand this phrase

If you have ticked (a), go straight to the next question
If you have ticked (b), you should also answer (c)

c) The kinds of thing I think of when I think of *having a comfortable life* include
Example 1 ....
Example 2 ....

Question 6 *somebody persuading someone else*

Tick either (a) or (b) below
a) I don't understand this phrase
b) I think I understand this phrase

If you have ticked (a), go straight to the next question
If you have ticked (b), you should also answer (c) and (d)

c) The kinds of thing I think of when I think of *somebody persuading someone else* include
Example 1 ....
Example 2 ....

Appendix I: Task 1, Part 2 (Experiment 2a & 2b)

When we speak of *"animal"*, the concept includes all the animals we know, such as horse, pig, dog, cat, mouse and so on. Similarly, the concept *"insect"* includes beetle, mosquito, fly and so on. However, when we use the word "animal" in everyday speech or writing, we can use it to refer to every animal or just those animals we are talking about when we speak.

(當我們說 **"動物"** 時，在我們的概念裡包含所有我們熟知的動物。例如馬、豬、狗、貓、老鼠等等。提到 **"昆蟲"** 時，則包含甲蟲、蚊子、蒼蠅等等。但是，當我們使用 **"動物"** 這個字眼時，它可以指每一種動物，或者只有指那些我們談論提及的動物。)

Example:
1. The animals we saw yesterday were elephants, lions, hippos and giraffes.
(昨天我們看到的動物是大象、獅子、河馬和長頸鹿。)

2. We kept cats, dogs, and birds.
(我們家養有貓、狗和小鳥。)

In the first example, *"animal"* includes *"elephants"*, *"lions"*, *"hippos"* and *"giraffes"*. However, in the second sentence, the concept *"animal"* includes *"cats"*, *"dogs"* and *"birds"*, though the category name *"animal"* is not specified.

（在第一句中， **"動物"** 包含了大象、獅子、河馬和長頸鹿。而在第二句中，則只包含貓、狗和小鳥，但是文句中沒有提到 **"動物"** 二字。）

3. As a full-time housewife, she has housework to do: sweeping the floor, vacuuming the carpet, doing the laundry, and cleaning the toilet.
(做一個全職的家庭主婦，她有家事要做：掃地，吸地毯，洗衣服，和洗廁所。)

Similarly in the example 3, the concept *"housework"* includes *"sweeping the floor"*, *"vacuuming the carpet"*, *"doing the laundry"*, and *"cleaning the toilet"*.
（同樣地，在第三句中， **"家事"** 包含掃地，吸地毯，洗衣服，洗廁所。）

Part 2: In the following questions, the words used may not be identical to the words in the text. In each question, first study the examples, then read the key, and after that circle A, B, C or D for each example.

Chapter 3

Question 1   *Derek's messy life*

Examples:

Example 1 *he was playing badly*	A B C D
Example 2 *he didn't have many friends*	A B C D
Example 3 *the person he thought he loved had disappeared from his life*	A B C D

Key:
A means 'I think I read this and it could be included in *Derek's messy life*'
B means 'I think I read this and it could not be included in *Derek's messy life*'
C means 'I don't think I read this and it could be included in *Derek's messy life*'
D means 'I don't think I read this and it could not be included in *Derek's messy life*'

Question 2   *ways of being a magician with the hat*

Examples:

Example 1 *running around the crowds that were listening*	A B C D
Example 2 *smiling at the crowds*	A B C D
Example 3 *laughing and joking*	A B C D
Example 4 *telling the crowds they were poor refugees*	A B C D

Key:
A means 'I think I read this and it could be included in *ways of being a magician with the hat*'
B means 'I think I read this and it could not be included in *ways of being a magician with the hat*'
C means 'I don't think I read this and it could be included in *ways of being a magician with the hat*'
D means 'I don't think I read this and it could not be included in *ways of being a magician with the hat*'

Question 3   *Derek's reasons for liking Rachel*

Examples:

Example 1 *she was quiet and gentle*	A B C D
Example 2 *she had light brown hair*	A B C D
Example 3 *she had pretty, brown eyes set in a round, pleasant face*	A B C D
Example 4 *when she smiled she looked like a happy child*	A B C D
Example 5 *you knew you could trust her*	A B C D
Example 6 *she was very easy to talk to*	A B C D

Key:
A means 'I think I read this and it could be included in *Derek's reasons for liking*

Rachel'

B means 'I think I read this and it could not be included in *Derek's reasons for liking Rachel'*

C means 'I don't think I read this and it could be included in *Derek's reasons for liking Rachel'*

D means 'I don't think I read this and it could not be included in *Derek's reasons for liking Rachel'*

Question 4 *Derek's hopes for the future*

Examples:
Example 1 *to make enough money to have a nice house*   A B C D
Example 2 *to travel a bit*   A B C D

Key:
A means 'I think I read this and it could be included in *Derek's hopes for the future'*

B means 'I think I read this and it could not be included in *Derek's hopes for the future'*

C means 'I don't think I read this and it could be included in *Derek's hopes for the future'*

D means 'I don't think I read this and it could not be included in *Derek's hopes for the future'*

Question 5 *things that weren't going well with Derek's studies*

Examples:
Example 1 *his viola teacher was ill*   A B C D
Example 2 *the man who took Derek's sick teacher's place didn't like him*   A B C D
Example 3 *Derek didn't like the man who took his sick teacher's place*   A B C D
Example 4 *Derek's piano playing wasn't getting much better*   A B C D

Key:
A means 'I think I read this and it could be included in *things that weren't going well with Derek's studies'*

B means 'I think I read this and it could not be included in *things that weren't going well with Derek's studies'*

C means 'I don't think I read this and it could be included in *things that weren't going well with Derek's studies'*

D means 'I don't think I read this and it could not be included in *things that weren't going well with Derek's studies'*

Question 6 *the life of the river*

Examples:
Example 1 *birds*   A B C D
Example 2 *the sounds of a great city*   A B C D
Example 3 *planes making their way westwards to the airport*   A B C D

274

Example 4 *rowers shouting orders at each other*   A B C D
Example 5 *a little police boat, its blue light flashing as it went as* A B C D
 fast as it could towards some emergency

Key:
A means 'I think I read this and it could be included in *the life of the river*'
B means 'I think I read this and it could not be included in *the life of the river*'
C means 'I don't think I read this and it could be included in *the life of the river*'
D means 'I don't think I read this and it could not be included in *the life of the river*'

Chapter 4

Question 1   *what happened later in Derek and Malgosia's relationship when they were unhappy*

Examples:

Example 1 *they started arguing*   A B C D
Example 2 *they didn't talk to each other*   A B C D
Example 3 *they were extremely polite*   A B C D

Key:
A means 'I think I read this and it could be included in *what happened later in Derek and Malgosia's relationship when they were unhappy'*
B means 'I think I read this and it could not be included in *what happened later in Derek and Malgosia's relationship when they were unhappy'*
C means 'I don't think I read this and it could be included in *what happened later in Derek and Malgosia's relationship when they were unhappy'*
D means 'I don't think I read this and it could not be included in *what happened later in Derek and Malgosia's relationship when they were unhappy'*

Question 2   *Tibor having a dream*

Examples:

Example 1 *having great wealth*   A B C D
Example 2 *having nothing to do except spending great wealth*   A B C D

Key:
A means 'I think I read this and it could be included in *Tibor having a dream'*
B means 'I think I read this and it could not be included in *Tibor having a dream'*
C means 'I don't think I read this and it could be included in *Tibor having a dream'*
D means 'I don't think I read this and it could not be included in *Tibor having a dream'*

Question 3   *all that had made Malgosia the person she was*

Examples:

Example 1 *her troubled childhood*   A B C D
Example 2 *that terrible business with Tibor*   A B C D

Key:
A means 'I think I read this and it could be included in *all that had made Malgosia the person she was'*
B means 'I think I read this and it could not be included in *all that had made Malgosia the person she was'*
C means 'I don't think I read this and it could be included in *all that had made Malgosia the person she was'*
D means 'I don't think I read this and it could not be included in *all that had made Malgosia the person she was'*

Question 4 *a long conversation*

Examples:
Example 1 *being silent*   **A B C D**
Example 2 *talking urgently in a low voice*   **A B C D**

Key:
A means 'I think I read this and it could be included in *a long conversation*'
B means 'I think I read this and it could not be included in *a long conversation*'
C means 'I don't think I read this and it could be included in *a long conversation*'
D means 'I don't think I read this and it could not be included in *a long conversation*'

Question 5 *Malgosia and Derek having comfortable lives*

Examples:
Example 1 *Malgosia getting work in musical shows*   **A B C D**
Example 2 *both of them doing some teaching too*   **A B C D**

Key:
A means 'I think I read this and it could be included in *Malgosia and Derek having comfortable lives*'
B means 'I think I read this and it could not be included in *Malgosia and Derek having comfortable lives*'
C means 'I don't think I read this and it could be included in *Malgosia and Derek having comfortable lives*'
D means 'I don't think I read this and it could not be included in *Malgosia and Derek having comfortable lives*'

Question 6 *Derek persuading Malgosia*

Examples:
Example 1 *she became Derek's lover*   **A B C D**
Example 2 *in the end she became Derek's wife*   **A B C D**

Key:
A means 'I think I read this and it could be included in *Derek persuading Malgosia*'
B means 'I think I read this and it could not be included in *Derek persuading Malgosia*'
C means 'I don't think I read this and it could be included in *Derek persuading Malgosia*'
D means 'I don't think I read this and it could not be included in *Derek persuading Malgosia*'

Appendix J: Task 2 (Experiment 3)

Recall the example you studied in task 1, 'We kept cats, dogs and birds'. *'Cats'*, *'dogs'* and *'birds'* belong to the same category, i.e. *'animal'*, though the category name *'animal'* is not specified.   Similarly, in the sentence 'My favourite fruits are bananas and mangos', *'bananas'* and *'mangos'* belong to the same category. That is, they are both fruit.   Unlike the previous example, the category name is present.

In each of the following questions two, three or four words or phrases or sentences belong to the same category. For each question you will be told how many items belong to the same category. You are required to place square brackets around each.

回憶你讀過的例句。 "我們家養有貓、狗和小鳥。" "貓" 、 "狗" 和 "小鳥" 在此句中是屬於同一類，雖然 "動物" 沒有寫出來，所以應該括出來。相同地， "我最喜歡的水果是香蕉和蘋果" ， "香蕉" 和 "蘋果" 在此句中是屬於同一類，也就是水果類，所以應該括出來。

下列的句子中，有些詞彙（可能是兩個字、三個字、或更長的字串組成）。其中，某些是屬於同一類。每一個問題將會告訴你有多少項目是屬於同類。請將這些同類的詞彙用方括號完整劃出來。是否同類，請用上下文決定。

Specimen Question 1
We kept cats, dogs, and birds. (3 items)
（我們家養有貓、狗和小鳥。）

Correct Answer
We kept [cats], [dogs], and [birds].
（答案：我們家養有【貓】、【狗】和【小鳥】。）

Specimen Question 2
My favourite fruits are bananas and mangos (2 items)
（我最喜歡的水果是香蕉和蘋果。）

Correct Answer
My favourite fruits are [bananas] and [mangos]
（答案：我最喜歡的水果是【香蕉】和【蘋果】。）

Now study the following questions in which two, three or four words or phrases or sentences belong to the same category. For each question you will be told how many items belong to the same category. You are required to place square brackets around each.

Question 1 (2 items)
I said, confused now and terribly worried.

Question 2 (2 items)
Oswaldo's nose says that would land us both in a lot of trouble, a great deal of trouble.

Question 3 (2 items)
It was quiet too. Occasionally a snatch of conversation from one of the houses came to us on the light wind, or the sound of a car in the distance.

Question 4 (2 items)
'She means that your Malgosia isn't here anymore, that's what,' Oswaldo said bluntly. 'She's gone.'

Question 5 (2 items)
I heard running feet and more shouting, but now the noises seemed to be getting further and further away.

Question 6 (2 items)
Strange noises.   Voices in the distance.   Footsteps going up and down somewhere near, echoing on a stone floor.

Question 7 (2 items)
That's why I probably behave a bit stupidly in dramatic situations, whether at Rosemary's front door or on a Brazilian hillside.

Question 8 (2 items)
But there was something else I had to ask.
'What about Malgosia?' I said, but nobody answered me. 'Come on,' I said. I was desperate to find out before I fell asleep again.

Question 9 (3 items)
.. who had added danger, smuggling and murder to the list of his charms.

Question 10 (3 items)
'Hey wait a minute!' Oswaldo's large voice bellowed down the phone line. 'Not so fast, OK? Just calm down.'

Question 11 (3 items)
We could hear a conversation, some shouting. And then suddenly I heard a voice I was sure I recognised, a female voice, Malgosia's voice.

Question 12 (3 items)

'Well this man, Reinaldo, he noticed your wife. The red hair, he said, she was beautiful. He couldn't take his eyes off her. She reminded him of an actress he was especially keen on.

Question 13 (4 items)
I prefer cities, lots of people, cars, noise, bars, all that kind of thing.

Question 14 (4 items)
But it didn't surprise him because this guy was always lucky. A man with powerful friends. Someone the police knew a lot about, but someone they'd never managed to pin anything on even though he was a real bad guy.

Question 15 (4 items)
I am not cut out for heroism or dramatic gestures. I'm a viola player. I love music, I love compromise. I'm even a bit boring.

Question 16 (4 items)
This time I was determined to open my eyes.
I managed to raise the corner of one eyelid with what seemed like a great effort. I was blinded by bright white light and shut it again. There was someone by the side of my bed. My bed? I was in a bed? What on earth was going on? I forced my eye open again, and managed to keep it open for a second longer.

Question 17 (4 items)
It was a hospital all right. There was a television on one wall, a large window with a view of the hills behind Rio, and light-green walls with nothing on them at all. I looked down at the bed. There were tubes coming out of my arm.

Question 18 (4 items)
'Malgosia!' I cried, 'Malgosia!' and ran down the hillside. They'd heard me now. Tibor turned round and looked to see where the noise was coming from. The other man stopped. Two more men ran out on to the terrace and looked up. They pulled out guns.
'Malgosia!' I cried again.

Appendix K: Task 3 (Experiment 4)

In each of the following questions please tick the version where in your opinion the words, phrases or sentences that belong in the same category have been underlined. Please tick only ONE version of each question.

下列各問題中，請圈選一個你認為劃線的各詞彙（可能是兩個字、三個字、或更長的字串組成）是屬於同類的。請只選一個答案。

Question 1

a) <u>Her hair was cut short</u>. She wore heavy black glasses. She watched me as I struggled in with my suitcase. <u>She didn't look welcoming at all</u>.

b) Her hair was cut short. She wore heavy black glasses. <u>She watched me as I struggled in with my suitcase</u>. <u>She didn't look welcoming at all</u>.

c) <u>Her hair was cut short</u>. She wore heavy black glasses. <u>She watched me as I struggled in with my suitcase</u>. She didn't look welcoming at all.

Question 2

a) Though <u>they still have to do more tests</u>. <u>They haven't been able to identify</u> what type it was yet.

b) Though <u>they still have to do more tests</u>. They haven't been able <u>to identify what type it was</u> yet.

c) Though <u>they still have to do more tests</u>. <u>They haven't been able to identify what type it was yet</u>.

Question 3

a) It was a bumpy journey, <u>with turbulence almost all the way</u>. Most of the time I half slept as the plane bounced through the skies.

b) <u>It was a bumpy journey</u>, with turbulence almost all the way. Most of the time I half slept as <u>the plane bounced through the skies.</u>

c) It was a bumpy journey, <u>with turbulence almost all the way</u>. Most of the time I half slept as <u>the plane bounced through the skies.</u>

Question 4

a) both typical symptoms of someone who had been poisoned by <u>some form of nerve gas</u>, <u>some kind of chemical agent</u>.

b) both <u>typical symptoms of someone who had been poisoned</u> by some form of nerve gas, <u>some kind of chemical agent</u>.

c) both typical symptoms of <u>someone who had been poisoned by some form of nerve gas</u>, <u>some kind of chemical agent</u>.

Question 5
a) Anja started up a low terrible moaning which seemed to go on and on. I wished she would just stop. <u>Even when her father made me go with him into the kitchen we could still hear it</u>, <u>the inhuman noise of someone who was going to feel guilty for ever</u>.

b) Anja started up <u>a low terrible moaning which seemed to go on and on</u>. I wished she would just stop. <u>Even when her father made me go with him into the kitchen we could still hear it</u>, the inhuman noise of someone who was going to feel guilty for ever.

c) Anja started up <u>a low terrible moaning which seemed to go on and on</u>. I wished she would just stop. Even when her father made me go with him into the kitchen we could still hear it, <u>the inhuman noise of someone who was going to feel guilty for ever</u>.

Question 6
a) Anja was <u>nowhere to be seen</u>. Nobody mentioned <u>her absence</u>.

b) Anja was <u>nowhere to be seen</u>. <u>Nobody mentioned her</u> absence.

c) Anja was <u>nowhere</u> to be seen. Nobody mentioned <u>her absence</u>.

Question 7
a) It took a bit of time <u>for his words to sink in</u>. It took me time <u>to realise that Malgosia couldn't hear me in there</u>. It took me time to realise that after all my journeying I had finally found her <u>and it wouldn't do me any good at all.</u>

b) It took a bit of time <u>for his words to sink in</u>. It took me time to realise that Malgosia couldn't hear me in there. It took me time <u>to realise that after all my journeying I had finally found her</u> and <u>it wouldn't do me any good at all</u>.

c) It took a bit of time <u>for his words to sink in</u>. It took me time <u>to realise that Malgosia couldn't hear me in there</u>. It took me time <u>to realise that after all my journeying I had finally found her and it wouldn't do me any good at all</u>.

Question 8

a) <u>'Poisoned?' I repeated</u>. I hadn't the slightest idea what he was talking about.

'<u>You didn't know? You didn't know about this</u>?'

b) 'Poisoned?' I repeated. <u>I hadn't the slightest idea what he was talking about</u>.

<u>You didn't know? You didn't know about this</u>?'

c) 'Poisoned?' I repeated. <u>I hadn't the slightest idea</u> what he was talking about.

'<u>You didn't know? You didn't know about this</u>?'

Question 9

a) every time he came here ... for years ... and then <u>she took him away from me</u>. <u>She</u> <u>took him back</u>.

b) <u>every time he came here</u> ... for years ... and then she took him away from me. <u>She</u> <u>took him back</u>.

c) <u>every time he came here</u> ... for years ... and then <u>she took him away from me</u>. She took him back.

Question 10

a) I was <u>too tired to feel anything</u>, <u>too shocked</u> to understand what had happened, <u>too</u> <u>alone to feel any warmth</u> from the people around me.

b) I was <u>too tired to feel anything</u>, <u>too shocked</u> to understand what had happened, <u>too</u> <u>alone to feel any warmth from the people around me</u>.

c) I was <u>too tired to feel anything</u>, <u>too shocked to understand what had happened</u>, <u>too</u> <u>alone to feel any warmth from the people around me</u>.

Appendix L: Task 4 (Experiment 5)

In the following questions, each item is accompanied by a rating scale. Based on your understanding of the text, ring the best answer for each item.

下列問題中，各項目旁邊列有評量表，請依照你對讀本的瞭解，判別各項目的程度，並圈選之。

For example:

Question 1: *It made me realise where Malgosia got her bright-eyed intelligence and her deep dark passion.*

In this question you are asked to decide how good a description of Malgosia's character the phrases *'her bright-eyed intelligence'* and *'her deep dark passion'* are. If you think *'her bright-eyed intelligence'* is a very good description of Malgosia's character, circle the answer **a very good**. If you think *'her bright-eyed intelligence'* is a good description, circle the answer **a good**. If you think *'her bright-eyed intelligence'* is an acceptable description, circle the answer **an acceptable**. If you think *'her bright-eyed intelligence'* is not a good description, circle the answer **not a good**.

Question 1

It made me realise where Malgosia got her bright-eyed intelligence and her deep dark passion.

Is

a) *her bright-eyed intelligence*

a very good	**a good**	**an acceptable**	**not a good**

b) *her deep dark passion*

a very good	**a good**	**an acceptable**	**not a good**

description of Malgosia's character?

Question 2

Now I was in a taxi, unhappy and confused, my head still bandaged, my mind trying to come to terms with all that had happened to me over the last week, my body tired

and hurting, the terrible plane smell in my hair and on my clothes.

Is

a) *my head still bandaged*

a very good	a good	an acceptable	not a good

b) *my mind trying to come to terms with all that had happened to me over the last week*

a very good	a good	an acceptable	not a good

c) *my body tired and hurting*

a very good	a good	an acceptable	not a good

d) *the terrible plane smell in my hair and on my clothes*

a very good	a good	an acceptable	not a good

reason for Derek to be unhappy and confused?

Question 3

they'd taken me to Krakow and Gdansk, driving over the long roads of this proud country with its unbeaten spirit, keeping its identity

Is

a) *its unbeaten spirit*

a very good	a good	an acceptable	not a good

b) *keeping its identity*

a very good	a good	an acceptable	not a good

reason for the country to be proud?

Question 4

The funeral was as bad as a funeral could be; grey rain coming at us almost horizontally from a grey cold sky, the voice of the priest at the graveside taken away from us by a bitter wind, the tears of the mourners invisible in the wet cold.

Would

a) *grey rain coming at us almost horizontally from a grey cold sky*

make a funeral	very bad indeed	very bad	bad	not bad

b) *the voice of the priest at the graveside taken away from us by a bitter wind*

make a funeral	very bad indeed	very bad	bad	not bad

c) *the tears of the mourners invisible in the wet cold*

make a funeral	very bad indeed	very bad	bad	not bad

Question 5

Malgosia's father managed to ask his questions, though. 'What kind of a husband do you call yourself?' was his first effort. 'What kind of a husband that lets this happen?

Is

a) *'What kind of a husband do you call yourself?'*

	a very good	a good	an acceptable	not a good

b) *'What kind of a husband that lets this happen?*

	a very good	a good	an acceptable	not a good

question for Malgosia's father to ask?

Question 6

'I feel terrible, all right?' I replied angrily. 'My wife's run off with someone else, I've been questioned by the police, I've travelled halfway round the world, I've been shot at and nearly killed. Now my wife's supposed to be very ill and I don't know why, or where she is, but I know it's pretty damn serious.

Is

a) *My wife's run off with someone else*

	a very good	a good	an acceptable	not a good

b) *I've been questioned by the police*

	a very good	a good	an acceptable	not a good

c) *I've travelled halfway round the world*

	a very good	a good	an acceptable	not a good

d) *I've been shot at and nearly killed*

	a very good	a good	an acceptable	not a good

e) *my wife's supposed to be very ill and I don't know why, or where she is, but I know it's pretty damn serious.*

         a very good   **a good**   **an acceptable**   **not a good**

reason for Derek to feel terrible?

Question 7

what Malgosia's father told me was that my wife had died, in the end, bleeding unstoppably, completely unable to breathe

Is

a) *bleeding unstoppably*

        a very typical   **a typical**   **a moderately typical**   **not a typical**

b) *being completely unable to breathe*

        a very typical   **a typical**   **a moderately typical**   **not a typical**

cause of death?

Question 8

there was a blank space in my brain where nothing moved, nothing was happening.

Is

a) *nothing moving*

        a very good   **a good**   **an acceptable**   **not a good**

b) *nothing happening*

        a very good   **a good**   **an acceptable**   **not a good**

description of a blank space in my brain?

Question 9

I'm just trying to absorb what you're telling me. I'm trying to come to terms with what's happened

Is

a) *absorbing what you're telling me*

something that would be

very difficult	**difficult**	**not too difficult**	**not difficult**

b) *come to terms with what's happened*

something that would be

very difficult	**difficult**	**not too difficult**	**not difficult**

to do?

Question 10

When the plane landed it taxied away towards the terminal building, but then, unexpectedly, it came to a stop.

Is

a) *it taxied away towards the terminal building*

a very good	**a good**	**an acceptable**	**not a good**

b) *it came to a stop*

a very good	**a good**	**an acceptable**	**not a good**

description of what happens when a plane lands?

Question 11

'And I'm not there.'

'Why not?' I shouted at her, 'why not?'

'Because I don't dare. Because I can't bear it. Because she was my best friend before Tibor.'

Is

a) *Because I don't dare*

a very good	**a good**	**an acceptable**	**not a good**

b) *Because I can't bear it*

a very good	**a good**	**an acceptable**	**not a good**

c) *Because she was my best friend before Tibor*

a very good	**a good**	**an acceptable**	**not a good**

reason for not being there?

References

Aitchison, J. (1987). *Words in the mind.* Oxford: Basil Blackwell.

Alcon, E. (1998). Input and input processing in second language acquisition. *IRAL, 36*(4), 343-352.

Alderson, J. C. (1984). Reading in a foreign language: A reading problem or a language problem? In J. C. Alderson & A. H. Urquhart (Eds.), *Reading in a foreign language* (pp. 1-27). Harlow, Essex: Longman.

Alderson, C., & Urquhart, A. H. (1984). *Reading in a foreign language.* Harlow, Essex: Longman.

Anderson, R. C., & Ortony, A. (1975). On putting apples into bottles--a problem of polysemy. *Cognitive Psychology, 7,* 167-180.

Anderson, R. C., Pichert, J. W., Goetz, E. T., Schallert, D. L., Stevens, K. V., & Trollip, S. R. (1976). Instantiation of general terms. *Journal of Verbal Learning and Verbal Behavior, 15,* 667-679.

Armstrong, S. L., Gleitman, L. R., & Gleitman, H. (1983). What some concepts might not be. *Cognition, 13,* 263-308.

Atkinson, R. C. (1975). Mnemotechnics in second-language learning. *American Psychologist, 30,* 821-828.

Atkinson, R. C., & Raugh, M. R. (1975). An application of the mnemonic keyword method to the acquisition of a Russian vocabulary. *Journal of Experimental Psychology: Human Learning and Memory, 104*(2), 126-133.

Avila, E., & Sadoski, M. (1996). Exploring new applications of the keyword method to acquiring English vocabulary. *Language Learning, 46*(3), 379-395.

Barclay, J. R., Bransford, J. D., Franks, J. J., McCarrell, N. S., & Nitsch, K. (1974). Comprehension and semantic flexibility. *Journal of Verbal Learning and*

Verbal Behavior, 13, 471-481.

Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition, 11*(3), 211-277.

Barsalou, L. W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 101-140). Cambridge: CUP.

Barsalou, L. W. (1991). Deriving categories to achieve goals. *Psychology of Learning and Motivation, 27,* 1-64.

Baxter, J. (1980). The dictionary and vocabulary behavior: A single word or a handful? *TESOL Quarterly, 14*(3), 325-336.

Bejoint, H. (1981). The foreign students' use of monolingual English dictionaries: A study of language needs and reference skills. *Applied Linguistics, 2*(3), 207-222.

Bejoint, H. (1987). The value of the dictionary in vocabulary acquisition. In A. P. Cowie (Ed.), *The dictionary and the language learner* (pp. 97-104). Tubingen: Max Niemeyer Verlag.

Bensoussan, M., & Laufer, B. (1984). Lexical guessing in context in EFL reading comprehension. *Journal of Research in Reading, 7*(1), 15-32.

Bialystok. (1983). Inferencing: Testing the "hypothesis-testing" hypothesis. In H. W. Seliger & M. H. Long (Eds.), *Classroom oriented research in second language acquisition* (pp. 104-123). Rowley, MA: Newbury House.

Bialystok, E., & Sharwood Smith, M. (1985). Interlanguage is not a state of mind: An evaluation of the construct for second-language acquisition. *Applied Linguistics, 6*(2), 101-118.

Biederman, I. (1972). Perceiving real-world scenes. *Science, 177,* 77-80.

Blau, E. K. (1982). The effect of syntax on readability for ESL students in Puerto Rico. *TESOL Quarterly, 16*(4), 517-528.

Bogaards, P. (2001). Lexical units and the learning of foreign language vocabulary. *Studies in Second Language Acquisition, 23*, 321-343.

Breen, M. P. (1985). Authenticity in the language classroom. *Applied Linguistics, 6*(1), 60-70.

Brown, R. (1958). How shall a thing be called? *Psychological Review, 65*(1), 14-21.

Brumfit, C., & Mitchell, R. (1990). The language classroom as a focus for research. In C. Brumfit & R. Mitchell (Eds.), *Research in the language classroom* (pp. 3-15). Oxford: Modern English Publications and the British Council.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. London: John Wiley & Sons.

Cameron, L. (1994). Organizing the world: Children's concepts and categories, and implications for the teaching of English. *ELT Journal, 48*, 28-39.

Carton, A. S. (1971). Inferencing: A process in using and learning language. In P. Pimsleur & T. Quinn (Eds.), *The psychology of second language learning* (pp. 45-58). Cambridge: CUP.

Cech, C. G., & Shoben, E. J. (1985). Context effects in symbolic magnitude comparisons. *Journal of Experimental Psychology: Learning, Memory and Cognition, 11*(2), 299-315.

Channell, J. (1981). Applying semantic theory to vocabulary teaching. *ELT Journal, 35*(2), 115-122.

Charles, C. M. (1995). *Introduction to education research*. New York: Longman.

Chaudron, C. (1985). Intake: On models and methods for discovering learners' processing of input. *Studies in Second Language Acquisition, 7*, 1-14.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: The MIT Press.

Chomsky, N. (1968). *Language and mind*. New York: Harcourt Brace & World.

Chomsky, N. (1976). *Reflections on language*. London: Temple Smith.

Chomsky, N. (1980). *Rules and representations*. Oxford: Basil Balackwell.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. New York: Praeger.

Clark, E. V. (1983). Meanings and concepts. In J. H. Flavell & E. M. Markman (Eds.), *Handbook of child psychology* (Vol. III, pp. 787-840). New York: John Wiley & Sons.

Coady, J. (1997). L2 vocabulary acquisition through extensive reading. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 225-237). Cambridge: CUP.

Coe, R. J. (2002). *It's the effect size, stupid: What effect size and why it is important*. Paper presented at the the British Educational Research Association annual conference, Exeter, 12-14 September, 2002.

Cohen, A. D. (1987). The use of verbal and imagery mnemonics in second-language vocabulary learning. *Studies in Second Language Acquisition, 9*, 43-62.

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8*, 240-247.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.

292

Crow, J. T., & Quigley, J. R. (1985). A semantic field approach to passive vocabulary acquisition for reading comprehension. *TESOL Quarterly, 19*(3), 497-513.

Cruse, D. A. (1977). The pragmatics of lexical specificity. *Journal of Linguistics, 13*, 153-164.

Cruse, D. A. (1986). *Lexical semantics*. Cambridge: CUP.

Davies, A. (1984). Simple, simplified and simplification: What is authentic? In C. Alderson & A. H. Urquhart (Eds.), *Reading in a foreign language* (pp. 181-195). Harlow, Essex: Longman.

Davis, J. N. (1989). Facilitating effects of marginal glosses on foreign language reading. *The Modern Language Journal, 73*(1), 41-48.

Day, R. R., & Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge: CUP.

de Bot, K., Paribakht, T. S., & Wesche, M. B. (1997). Toward a lexical processing model for the study of second language vocabulary acquisition. *Studies in Second Language Acquisition, 19*, 309-329.

Dirven, R. (1989). Cognitive linguistics and pedagogic grammar. In G. Graustein & G. Leitner (Eds.), *Reference grammars and modern linguistic theory* (pp. 56-75). Tubingen: Niemeyer.

Dirven, R., Niemeier, S., & Putz, M. (2001). Introduction. In R. Dirven & S. Niemeier & M. Putz (Eds.), *Applied cognitive linguistics II: Language pedagogy* (pp. xiii-xxv). Berlin: Mouton de Gruyter.

Donoghue, D., & Mitchell, B. (1992). Parataxis and hypotaxis: A review of some terms used for old English syntax. *Neuphilologische Mitteilungern, 93*(2), 163-183.

Dubin, F. (1989). The odd couple: Reading and vocabulary. *ELT Journal, 43*(4), 283-287.

Dubin, F., & Olshtain, E. (1993). Predicting word meanings from contextual clues: Evidence for ll readers. In M. Huckin & M. Haynes & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 181-201). Norwood, NJ: Ablex.

Ellis, R. (1994). *The study of second language acquisition*. Oxford: OUP.

Ellis, R. (1995). Modified oral input and the acquisition of word meanings. *Applied Linguistics, 16*(4), 409-441.

Erbaugh, M. S. (1986). Taking stock: The development of Chinese noun classifiers historically and in young children. In C. Craig (Ed.), *Noun classes and categorization* (pp. 399-436). Amsterdam: John Benjamins.

Fesmire, S. A. (1994). What is "cognitive" about cognitive linguistics? *Metaphor and Symbolic Activity, 9*(2), 149-154.

Fraser, C. A. (1998). The role of consulting a dictionary in reading and vocabulary learning. *Canadian Journal of Applied Linguistics, 2*(1-2), 73-89.

Fraser, C. A. (1999). Lexical processing strategy use and vocabulary learning through reading. *Studies in Second Language Acquisition, 21*(2), 225-241.

Gardner, R. (1980). On the validity of affective variables in second language acquisition: Conceptual, contextual, and statistical considerations. *Language Learning, 30*, 255-270.

Gardner, R. (1988). The socio-educational model of second language learning: Assumptions, findings and issues. *Language Learning, 38*, 101-126.

Gass, S. (1979). Language transfer and universal grammatical relations. *Language Learning, 29*(2), 327-344.

Gass, S. (1988). Integrating research areas: A framework for second language studies. *Applied Linguistics, 9*(2), 198-217.

Gass, S. (1989). Second language vocabulary acquisition. *Annual Review of Applied Linguistics, 9*(1), 92-106.

Gass, S. (1995). Learning and teaching: The necessary intersection. In F. R. Eckman (Ed.), *Second language acquisition theory and pedagogy* (pp. 3-20). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gass, S. (1997). *Input, interaction, and the second language learner.* Mahwah, NJ: Lawrence Erlbaum Associates.

Gass, S. (1999). Discussion: Incidental vocabulary learning. *Studies in Second Language Acquisition, 21*, 319-333.

Gass, S., & Varonis, E. M. (1994). Input, interaction, and second language production. *Studies in Second Language Acquisition, 16*, 283-302.

Geach, P., & Black, M. (1980). *Translations from the philosophical writings of Gottlob Frege.* Oxford: Basil Blackwell.

Geeraerts, D. (1989). Prospects and problems of prototype theory: Introduction. *Linguistics, 27*, 587-612.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition, 23*, 183-209.

Gentner, D., & Namy, L. L. (1999). Comparison in the development of categories. *Cognitive Development, 14*, 487-513.

Gibbs, R. W. (1995). What's cognitive about cognitive linguistics?, *Cognitive linguistics in the redwoods* (pp. 27-53). Berlin: Mouton de Gruyter.

Gilliland, J. (1972). *Readability.* London: Hodder and Stoughton Educational.

Gonzalez, O. (1999). Building vocabulary: Dictionary consultation and the ESL

student. *Journal of Adolescent & Adult Literacy, 43*(3), 264-270.

Goodman, K. S. (1976). Reading: A psycholinguistic guessing game. In H. Singer & R. Ruddell (Eds.), *Theoretical models and processes of reading* (pp. 259-271). Newark, Del.: International Reading Association.

Grabe, W., & Stoller, F. L. (1997). Reading and vocabulary development in a second language: A case study. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 98-122). Cambridge: CUP.

Grace, C. A. (1998). Retention of word meanings inferred from context and sentence-level translations: Implications for the design of beginning-level CALL software. *The Modern Language Journal, 82*(4), 533-544.

Haastrup, K. (1991). *Lexical inferencing procedures or talking about words: Receptive procedures in foreign language learning with special reference to English*. Tubingen: Gunter Narr Verlag.

Hafiz, F. M., & Tudor, I. (1989). Extensive reading and the development of language skills. *ELT Journal, 43*(1), 4-13.

Hampton, J. (1993). Prototype models of concept representation. In I. V. Mechelen & J. Hampton & R. S. Michalski & P. Theuns (Eds.), *Categories and concepts: Theoretical views and inductive data analysis* (pp. 67-95). London: Academic Press.

Harmer, J. (1999). *Trumpet voluntary*. Cambridge: CUP.

Hartmann, R. R. K. (1987). Four perspectives on dictionary use: A critical review of research methods. In A. P. Cowie (Ed.), *The dictionary and the language learner: Papers from the EURALEX seminar at the University of Leeds, 1-3 April 1985* (pp. 11-28). Tubingen: Max Niemeyer Verlag.

Haynes, M. (1993). Patterns and perils of guessing in second language reading. In M.

Huckin & M. Haynes & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 46-64). Norwood, NJ: Ablex.

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition, 21*(2), 303-317.

Hill, D. R. (1997). Survey review: Graded readers. *ELT Journal, 51*(1), 57-79.

Hill, D. R. (2001). Graded readers. *ELT Journal, 55*(3), 300-324.

Hill, D. R., & Thomas, H. R. (1988). Survey review: Graded readers (part 1). *ELT Journal, 42*(1), 44-52.

Hogaboam, T. W., & Pellegrino, J. W. (1978). Hunting for individual differences in cognitive processes: Verbal ability and semantic processing of pictures and words. *Memory & Cognition, 6*(2), 189-193.

Horsfall, P. (1997). Dictionary skills in MFL 11-16. *Language Learning Journal, 15*, 3-9.

Huckin, M., & Bloch, J. (1993). Strategies for inferring word meaning in context: A cognitive model. In M. Huckin & M. Haynes & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 153-178). Norwood, NJ: Ablex.

Huckin, T., & Coady, J. (1999). Incidental vocabulary acquisition in a second language. *Studies in Second Language Acquisition, 21*(2), 181-193.

Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In P. J. L. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 113-125). Basingstok, Hants: Macmillan.

Hulstijn, J. H. (1993). When do foreign-language readers look up the meaning of unfamiliar words? The influence of task and learner variables. *The Modern*

Language Journal, 77(2), 139-147.

Hulstijn, J. (1997a). Mnemonic methods in foreign language vocabulary learning: Theoretical considerations and pedagogical implications. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 203-224). Cambridge: CUP.

Hulstijn, J. (1997b). Second language acquisition research in the laboratory: Possibilities and limitations. *Studies in Second Language Acquisition, 19*(1), 131-143.

Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal, 80*(3), 327-339.

Jacobs, G. M., Dufon, P., & Hong, F. C. (1994). L1 and L2 vocabulary glosses in L2 reading passages: Their effectiveness for increasing comprehension and vocabulary knowledge. *Journal of Research in Reading, 17*, 19-28.

Jakobson, R. (1960). Linguistics and poetics. In T. A. Sebeok (Ed.), *Style in language* (pp. 350-377). New York: Technology Press of Massachusetts Institute of Technology and John Wiley & Sons.

Kellerman, E. (1979). Transfer and non-transfer: where are we now. *Studies in Second Language Acquisition, 2*, 37-57.

Kellerman, E. (1983). Now you see it, now you don't. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 112-134). Rowley, MA: Newbury House.

Kembo, J. A. (2001). Testing of inferencing behaviour in a second language. *International Journal of Bilingual Education and Bilingualism, 4*(2), 77-96.

King, J. R. (1984). Levels of categorization and sight word acquisition. *Reading Psychology, 5*, 129-138.

Klapper, J. (1993). Practicable skills and practical constraints in FL reading. *Language Learning Journal, 7*, 50-54.

Knight, S. (1994). Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal abilities. *The Modern Language Journal, 78*(3), 285-299.

Koda, K. (1997). Orthographic knowledge in L2 lexical processing: A cross-linguistic perspective. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 35-52). Cambridge: CUP.

Krashen, S. (1981). *Second language acquisition and second language learning.* Oxford: Pergamon.

Krashen, S. (1982). *Principles and practice in second language acquisition.* Oxford: Pergamon.

Krashen, S. (1985). *The input hypothesis: Issues and implications.* London: Longman.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal, 73*(4), 440-464.

Kruse, H., Pankhurst, J., & Sharwood Smith, M. (1987). A multiple word association probe in second language acquisition research. *Studies in Second Language Acquisition, 9*, 141-154.

Labov, W. (1973). The boundaries of words and their meanings. In C.-J. N. Bailey & R. W. Shuy (Eds.), *New ways of analyzing variation in English* (pp. 340-373). Washington, D.C.: Georgetown University Press.

Lakoff, G. (1986). Classifiers as a reflection of mind. In C. Craig (Ed.), *Nouns classes and categorization* (pp. 13-51). Amsterdam: John Benjamins.

Lakoff, G. (1987). *Women, fire, and dangerous things*. Chicago: The University of Chicago Press.

Lakoff, G., & Johnson, D. M. (1999). *Philosophy in the flesh*. New York: Basic Books.

Langacker, R. W. (1997). The contextual basis of cognitive semantics. In J. Nuyts & E. Pederson (Eds.), *Language and conceptualization* (pp. 229-252). Cambridge: CUP.

Larsen-Freeman, D. (1996). The changing nature of second language classroom research. In J. Schachter & S. Gass (Eds.), *Second language classroom research: Issues and opportunities* (pp. 157-170). Mahwah, NJ: Lawrence Erlbaum Associates.

Lassaline, M. E., Wisniewski, E. J., & Medin, D. L. (1992). The basic level in artificial and natural categories: Are all basic levels created equal? In B. Burns (Ed.), *Percepts, concepts and categories: The representation and processing of information* (pp. 327-378). Amsterdam: Elsevier.

Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. J. L. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 126-132). Basingstok, Hants: Macmillan.

Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 20-34). Cambridge: CUP.

Laufer, B., & Kimmel, M. (1997). Bilingualised dictionaries: How learners really

use them. *System, 25*(3), 361-369.

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics, 16*(3), 307-322.

Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing, 16*(1), 33-51.

Laufer, B., & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning, 48*(3), 365-391.

Lawson, M., & Hogben, D. (1996). The vocabulary-learning strategies of foreign-language students. *Language Learning, 46*, 101-135.

Leow, R. P. (1993). To simplify or not to simplify. *Studies in Second Language Acquisition, 15*, 333-355.

Levin, S. R. (1970). Coupling in Shakespearean sonnet. In D. C. Freeman (Ed.), *Linguistics and literary style* (pp. 197-205). New York: Holt, Rinehart and Winston.

Lewis, M. (1993). *The lexical approach.* Hove: Language Teaching Publications.

Li, X. (1988). Effects of contextual cues on inferring and remembering meanings of new words. *Applied Linguistics, 9*(4), 402-413.

Lin, E. L., Murphy, G. L., & Shoben, E. J. (1997). The effects of prior processing episodes on basic-level superiority. *The Quarterly Journal of Experimental Psychology, 50A*(1), 25-48.

Long, M. H. (1983). Inside the "black box": Methodological issues in classroom research on language learning. In H. W. Seliger & M. H. Long (Eds.), *Classroom oriented research in second language acquisition* (pp. 3-35). Rowley, MA: Newbury House.

Long, M. H. (1990). Second language classroom research and teacher education. In C. Brumfit & R. Mitchell (Eds.), *Research in the language classroom* (pp. 161-170). Oxford: Modern English Publications and the British Council.

Luppescu, S., & Day, R. R. (1993). Reading, dictionaries, and vocabulary learning. *Language Learning, 43*(2), 263-287.

MacIntyre, P. D., & Gardner, R. C. (1991). Methods and results in the study of anxiety and language learning : A review of the literature. *Language Learning, 41*(1), 85-117.

MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning, 44*(2), 283-305.

Markman, E. M. (1987). How children constrain the possible meanings of words. In U. Neisser (Ed.), *Concepts and conceptual development: Ecologicl and intellectual factors in categorization* (pp. 255-287). Cambridge: CUP.

Melka, F. (1997). Receptive vs. productive aspects of vocabulary. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 84-102). Cambridge: CUP.

McCloskey, M. E., & Glucksberg, S. (1978). Natural categories: Well defined or fuzzy sets? *Memory & Cognition, 6*(4), 462-472.

McDaniel, M. A., & Pressley, M. (1984). Putting the keyword method in context. *Journal of Educational Psychology, 76*(4), 598-609.

McKoon, G., & Ratcliff, R. (1988). Contextually relevant aspects of meaning. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14*(2), 331-343.

McKoon, G., & Rafcliff, R. (1989). Inferences about contextually defined categories.

Journal of Experimental Psychology: Learning, Memory, and Cognition, 15(6), 1134-1146.

McMillan, J. H., & Schumacher, S. (1997). *Research in education: A conceptual introduction*. New York: Longman.

Meara, P. (1984). The study of lexis in interlanguage. In A. Davies & C. Criper & A. P. R. Howatt (Eds.), *Interlanguage* (pp. 225-235). Endinburgh: Edinburgh University Press.

Meara, P. (1990). A note on passive vocabulary. *Second Language Research, 6*, 151-154.

Meara, P. (1993). The bilingual lexicon and the teaching of vocabulary. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 279-297). Amsterdam: John Benjamins.

Meara, P., & Jones, G. (1990). *Eurocentres vocabulary size tests 10KA*. Zurich: Eurocentres Learning Service.

Medin, D. L., & Barsalou, L. W. (1987). Categorization processes and categorical perception. In S. Harnad (Ed.), *Categorical perception: The ground of cognition* (pp. 455-490). Cambridge: CUP.

Mervis, C. B. (1980). Category structure and the development of categorization. In R. J. Spiro & B. C. Bruce & W. F. Brewer (Eds.), *The psychology of reading* (pp. 279-307). Hillsdale, NJ: Lawrence Erlbaum Associates.

Mervis, C. B., & Crisafi, M. A. (1982). Order of acquisition of subordinate-, basic-, and superordinate-level categories. *Child Development, 53*, 258-266.

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1975). Loci of contextual effects on visual word recognition. In P. M. A. Rabbit & S. Dornic (Eds.), *Attention and performance* (Vol. V, pp. 98-118). London: Academic Press.

Mondria, J.-A., & Wit-de Boer, M. (1991). The effects of contextual richness on the guessability and the retention of words in a foreign language. *Applied Linguistics, 12*(3), 249-267.

Moore, J. C., & Surber, J. R. (1992). Effects of context and keyword methods on second language vocabulary acquisition. *Contemporary Educational Psychology, 17,* 286-292.

Morris, M. W., & Murphy, G. L. (1990). Converging operations on a basic level in event taxonomies. *Memory & Cognition,* 407-418.

Mukarovsky, J. (1970). Standard language and poetic language. In D. C. Freeman (Ed.), *Linguistics and literary style* (pp. 40-56). New York: Holt, Rinehart and Winston.

Murphy, G. L. (1991). Meaning and concepts. In P. J. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 11-35). Hillsdale, NJ: Lawrence Erlbaum Associates.

Murphy, G. L., & Brownell, H. H. (1985). Category differentiation in object recognition: Typicality constraints on the basic category advantage. *Journal of Experimental Psychology: Learning, Memory and Cognition, 11,* 70-84.

Murphy, G. L., & Lassaline, M. E. (1997). Hierarchical structure in concepts and the basic level of categorization. In K. Lamberts & D. Shanks (Eds.), *Knowledge, concepts and categories* (pp. 93-131). East Sussex: Psychology Press.

Murphy, G. L., & Smith, E. E. (1982). Basic-level superiority in picture categorization. *Journal of Verbal Learning and Verbal Behavior, 21,* 1-20.

Murphy, G. L., & Wisniewski, E. J. (1989). Categorizing objects in isolation and in scenes: What a superordinate is good for. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(4), 572-586.

Na, L., & Nation, I. S. P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal, 16*, 35-42.

Nagy, W. E., & Herman, P. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 19-35). Hillsdale, NJ: Lawrence Erlbaum Associates.

Nagy, W. E., Herman, P., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly, 20*, 233-253.

Nation, I. S. P. (1990). *Teaching and learning vocabulary.* New York: Heinle & Heinle.

Nation, P. (2000). Learning vocabulary in lexical sets: Dangers and guidelines. *TESOL Journal, 9*(2), 6-10.

Nation, P., & Coady, J. (1988). Vocabulary and reading. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 97-110). London: Longman.

Nation, P., & Wang, M.-t., Karen. (1999). Graded readers and vocabulary. *Reading in a Foreign Language, 12*(2), 355-380.

Nation, P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6-19). Cambridge: CUP.

Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and languages teaching.* Oxford: OUP.

Nurweni, A., & Read, J. (1999). The English vocabulary knowledge of Indonesian university students. *English for Specific Purposes, 18*(2), 161-175.

Nuttall, C. (1982). *Teaching reading skills in a foreign language.* London:

Heinemann.

Nuttall, C. (1996). *Teaching reading skills in a foreign language* (2nd ed.). Oxford: Hinmann.

Odlin. (1989). *Language transfer.* Cambridge: CUP.

Oxford, R. L., & Scarcella, R. C. (1994). Second language vocabulary learning among adults: State of the art in vocabulary instruction. *System, 22*(2), 231-243.

Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition, 3*(5), 519-526.

Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 174-200). Cambridge: CUP.

Paribakht, T. S., & Wesche, M. (1999). Reading and "incidental" L2 vocabulary acquisition: An introspective study of lexical inferencing. *Studies in Second Language Acquisition, 21*(2), 195-224.

Parry, K. (1997). Vocabulary and comprehension: Two portraits. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 55-68). Cambridge: CUP.

Pellegrino, J. W., Rosinski, R. R., Chiesi, H. L., & Seiegel, A. (1977). Picture-word differences in decision latency: An analysis of single and dual memory models. *Memory & Cognition, 5*(4), 383-396.

Pica, T. (1991). Input as a theoretical and research construct: From Corder's original definition to current views. *IRAL, 29*(3), 185-196.

Pica, T. (1994). Questions from the language classroom: Research perspecitves.

TESOL Quarterly, 28(1), 49-79.

Pressley, M., Levin, J. R., Hall, J. W., Miller, G. E., & Berry, J. K. (1980). The keyword method and foreign word acquisition. *Journal of Experimental Psychology: Human Learning and Memory, 6*(2), 163-173.

Pressley, M., Levin, J. R., & McDaniel, M. A. (1987). Remembering versus inferring what a word means: Mnemonic and contextual approaches. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 107-127). Hillsdale, NJ: Lawrence Erlbaum Associates.

Pressley, M., Levin, J. R., Nakamura, G. V., Hope, D. J., Bispo, J. G., & Toye, A. R. (1980). The keyword method of foreign vocabulary learning: an investigation of its generalizability. *Journal of Applied Psychology, 65*(6), 635-642.

Prince, P. (1996). Second language vocabulary learning: The role of context versus translations as a function of proficiency. *The Modern Language Journal, 80*(4), 478-493.

Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *The Canadian Modern Language Review, 56*(2), 282-307.

Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing, 10*, 355-371.

Read, J. (2000). *Assessing vocabulary*. Cambridge: CUP.

Richards, J. C. (1976). The role of vocabulary teaching. *TESOL Quarterly, 10*(1), 77-89.

Riegel, K. F., Ramsey, R. M., & Riegel, R. M. (1967). A comparison of the first and second languages of American and Spanish students. *Journal of Verbal Learning and Verbal Behavior, 6*, 536-544.

Rifkin, A. (1985). Evidence for a basic level in event taxonomies. *Memory & Cognition, 13,* 538-556.

Rodriguez, M., & Sadoski, M. (2000). Effects of rote, context, keyword, and context/keyword methods on retention of vocabulary in EFL classroom. *Language Learning, 50*(2), 385-412.

Rosch, E. (1973a). Natural categories. *Cognitive Psychology, 4,* 328-350.

Rosch, E. (1973b). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 111-144). New York: Academic Press.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104*(3), 192-233.

Rosch, E. (1977). Human categorization. In N. Warren (Ed.), *Studies in cross-culture psychology* (Vol. 1, pp. 1-49). London: Academic Press.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27-48). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rosch, E., & Mervis, C., B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 6,* 573-605.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8,* 382-439.

Rossiter, M. J. (2001). The challenges of classroom-based SLA research. *Applied Language Learning, 12*(1), 31-44.

Roth, E. M., & Shoben, E. J. (1983). The effect of context on the structure of categories. *Cognitive Psychology, 15,* 346-378.

Rounds, P. L. (1996). The classroom-based researcher as fieldworker: Strangers in

a strange land. In j. Schachter & S. Gass (Eds.), *Second language classroom research: Issues and opportunities* (pp. 45-59). Mahwah, NJ: Lawrence Erlbaum Associates.

Rudzka-Ostyn, B. (1993). Introduction. In R. A. Geiger & B. Rudzka-Ostyn (Eds.), *Conceptualizations and mental processing in language* (pp. 1-20). Berlin: Mouton de Gruyter.

Rynell, A. (1952). *Parataxis and hypotaxis as a criterion of syntax and style.* Lund: C.W.K. Gleerup.

Saeed, J. I. (1997). *Semantics.* Oxford: Blackwell.

Saleemi, A. P. (1989). Inputs for L2 acquisition. *IRAL, 27*(2), 173-191.

Schachter, J. (1993). Second language acquisition: Perceptions and possibilities. *Second Language Research, 9*(2), 173-187.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158.

Schmitt, N. (2000). *Vocabulary in language teaching.* Cambridge: CUP.

Schmitt, N., & Carter, R. (2000). The lexical advantages of narrow reading for second language learners. *TESOL Journal, 9*(1), 4-9.

Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge framework: Word associations and verbal suffixes. *Studies in Second Language Acquisition, 19*, 17-36.

Schofield, P. (1982). Using the English dictionary for comprehension. *TESOL Quarterly, 16*(2), 185-194.

Scholfield, P. (1997). Vocabulary reference works in foreign language learning. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 279-302). Cambridge: CUP.

Schoonen, R., Hulstijn, J. H., & Bossers, B. (1998). Metacognitive and language-specific knowledge in native and foreign language reading comprehension: An empirical study among Dutch students in grades 6, 8 and 10. *Language Learning, 48*(1), 71-106.

Schouten-van Parrenen, C. (1989). Vocabulary through reading: Which conditions should be met when presenting words in texts? *AILA Review, 6,* 75-85.

Schuberth, R. E., & Eimas, P. D. (1977). Effects of context in the classification of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 27-36.

Schvaneveldt, R. W., Meyer, D. E., & Becker, C. A. (1976). Lexical ambiguity, semantic context, and visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 2*(2), 243-256.

Seliger, H. W., & Long, M. H. (1983). Introduction: What is classroom oriented research? In H. W. Seliger & M. H. Long (Eds.), *Classroom oriented research in second language acquisition* (pp. vii-xi). Rowley, MA: Newbury House.

Sharwood Smith, M. (1991). Speaking to many minds: On the relevance of different types of language information for the L2 learner. *Second Language Research, 7*(2), 118-132.

Sharwood Smith, M. (1993). Input enhancement in instructed SLA. *Studies in Second Language Acquisition, 15,* 165-179.

Smith, E. E. (1990). Categorization. In D. N. Osherson & E. E. Smith (Eds.), *An invitation to cognitive science* (Vol. 3. Thinking, pp. 33-53). Cambridge, MA: MIT Press.

Smith, E. E. (1995). Concepts and categorization. In E. E. Smith & D. N. Osherson (Eds.), *An invitation to cognitive science: Thinking* (2nd ed., pp. 3-33).

Cambridge, MA: MIT.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts.* Cambridge, MA: Harvard University Press.

Snell-Hornby, M. (1984). *The bilingual dictionary-help or hindrance?* Paper presented at the International Conference on Lexicography at Exeter, 9-12 September 1983, Exeter, England.

Snell-Hornby, M. (1987). Towards a learners' bilingual dictionary. In A. P. Cowie (Ed.), *The dictionary and the language learner* (pp. 159-170). Tubingen: Max Niemeyer Verlag.

Sokmen, A. J. (1997). Current trends in teaching second language vocabulary. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 237-257). Cambridge: CUP.

Spada, N., Ranta, L., & Lightbown, P. M. (1996). Working with teachers in second language acquisition research. In J. Schachter & S. Gass (Eds.), *Second language classroom research: Issues and opportunities* (pp. 31-44). Mahwah, NJ: Lawrence Erlbaum Associates.

Sternberg, R. J. (1987). Most vocabulary is learned from context. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 89-105). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sternberg, R. J., & Powell, J. S. (1983). Comprehending verbal comprehension. *American Psychologist*, 878-893.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H.G. Widdowson* (pp. 125-145). Oxford: OUP.

Swain, M. (1997). The influence of the mother tongue on second language vocabulary acquisition and use. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 156-180). Cambridge: CUP.

Swain, M., & Lapkin, S. (1994). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics, 16*, 371-391.

Tai, J. H. Y., & Chao, F.-y. (1994). A semantic study of the classifier *zhang*. *Journal of the Chinese Language Teachers Association, 29*(3), 67-78.

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology, 23*, 457-482.

Taylor, J. R. (1995). *Linguistic categorization: Prototypes in linguistic theory* (2nd ed.). Oxford: Clarendon Press.

Tinkham, T. (1997). The effects of semantic and thematic clustering on the learning of second language vocabulary. *Second Language Research, 13*(2), 138-163.

Tomaszczyk, J. (1983). On bilingual dictionaries. In R. R. K. Hartmann (Ed.), *Lexicography: Principles and practice* (pp. 41-50). London: Academic Press.

Tudor, I., & Hafiz, F. M. (1989). From input to intake-the effect of simplified readers on ESL development. *Journal of Reading, 32*(8), 688-693.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84*(4), 327-352.

Tversky, A., & Gati, I. (1978). Studies of similarity. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 79-98). Hillsdale, NJ: Lawrence

Erlbaum Associates.

Tversky, B., & Hemenway, K. (1983). Categories of environmental scenes. *Cognitive Psychology, 15,* 121-149.

Twaddell, F. (1973). Vocabulary expansion in the TESOL classroom. *TESOL Quarterly, 7*(1), 61-78.

Ungerer, F., & Schmid, H.-J. (1996). *An introduction to cognitive linguistics.* New York: Addison Wesley Longman.

van Lier, L. (1988). *The classroom and the language learner: Ethnography and second-language classroom research.* New York: Longman.

van Parreren, C. F., & Schouten-van Patteren, M. C. (1982). Contextual guessing: A trainable reader strategy. *System, 9*(3), 235-241.

Vespoor, M., & Winitz, H. (1997). Assessment of the lexical-input approach for intermediate language learners. *IRAL, 35*(1), 61-75.

Vu, H., Kellas, G., & Paul, S. T. (1998). Sources of sentence constraint on lexical ambiguity resolution. *Memory & Cognition, 26*(5), 979-1001.

Vygotskii, L. S. (1986). *Thought and language (translation newly revised and edited by Alex Kozulin).* Cambridge, MA: The MIT Press.

Waring, R. (1997). The negative effects of learning words in semantic sets: A replication. *System, 25*(2), 261-274.

Watanabe, Y. (1997). Input, intake, and retention. *Studies in Second Language Acquisition, 19,* 287-307.

Wertsch, J. V., & Tulviste, P. (1996). L.S. Vygotsky and comtemporary development psychology. In H. Daniels (Ed.), *An introduction to Vytotsky* (pp. 53-74). London: Routledge.

Wesche, M., & Paribakht, T. S. (1996). Assessing second language vocabulary

knowledge: Depth versus breadth. *Canadian Modern Language Review, 53*(1), 13-40.

Wesche, M., & Paribakht, T. S. (2000). Reading-based exercises instruction in a second language: An introspective study. *The Modern Language Journal, 84*(2), 196-213.

White, L. (1987). Against comprehensible input: The input hypothesis and the development of second-language competence. *Applied Linguistics, 6*(2), 95-110.

Widdowson, H. G. (1978). *Teaching language as communication*. Oxford: OUP.

William, J. (1999). Learner-generated attention to form. *Language Learning, 49*(4), 583-625.

Williams, E. (1983). Communicative reading. In K. Johnson & D. Porter (Eds.), *Perspectives in communicative language teaching* (pp. 171-188). London: Academic Press.

Wisniewski, E. J., & Murphy, G. L. (1989). Superordinate and basic category names in discourse: A textual analysis. *Discourse Processes, 12*, 245-261.

Wittgenstein, L. (1953). *Philosophical Investigations. Translated by G. E. M. Anscombe*. Oxford: Blackwll.

Wodinsky, M., & Nation, P. (1988). Learning for graded readers. *Reading in a Foreign Language, 5*(1), 155-161.

Yang, L. R., & Givon, T. (1997). Benefits and drawbacks of controlled laboratory studies of second language acquisition. *Studies in Second Language Acquisition, 19*, 173-193.

Ying, H. (1995). What sort of input is needed for intake? *IRAL, 33*, 173-194.

Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring vocabulary through reading:

Effects of frequency and contextual richness. *The Canadian Modern Language Review, 57*(4), 541-572.

Zimmerman, C. B. (1997). Do reading and interactive vocabulary instruction make a difference? An empirical study. *TESOL Quarterly, 3*(1), 121-140.