

**INVESTIGATING THE EFFECTS OF LEARNING TASKS ON
VOCABULARY KNOWLEDGE**

by

Stuart Alexander Webb

A thesis
submitted to the Victoria University of Wellington
in fulfillment of the
requirements for the degree of
Doctor of Philosophy
in Applied Linguistics

Victoria University of Wellington

2002

ABSTRACT

This thesis looks at whether different kinds of vocabulary learning tasks result in different types of word knowledge. In almost every study that has investigated the effects of tasks on vocabulary learning; the only aspect of word knowledge that was tested was meaning and form. Since researchers agree that knowing a word involves much more than knowing its meaning and form, prior research may have measured partial knowledge of only one of several aspects of knowledge. In order to determine the effects of vocabulary learning tasks, several aspects of knowledge should be tested. The experiments in this thesis investigated how vocabulary learning tasks affect both receptive and productive knowledge of five aspects of word knowledge: orthography, association, syntax, meaning and form, and grammatical functions. In the first of six experiments, the effects of incidental learning from reading and explicit learning from word pairs on word knowledge were compared. The results indicated that gains in knowledge tend to increase as the number of repetitions increases; however, partial gains from an informative context may be reduced or eliminated if followed by a less informative context. The results also showed that learning from word pairs contributed to surprisingly greater gains in all of the aspects. In the second experiment, two tasks (learning from glossed sentences, and learning from word pairs) were compared to determine the effects of context and synonymy on vocabulary knowledge. It was found that the subjects gained greater knowledge of unknown words that had high frequency synonyms than for those with less frequent synonyms. The results also indicated that a single context may have little effect on acquisition. In the third, fourth and fifth experiments, the effects of receptive and productive learning tasks on vocabulary knowledge were examined. The results indicated that productive learning from word pairs may be more effective at developing productive knowledge while receptive learning from word pairs may be more effective at increasing receptive knowledge. The sixth experiment investigated the effects of receptive and productive learning from word pairs on communication. It was found that the receptive task may be superior in improving comprehension, and the productive task may be better suited to facilitating

writing. Taken as a whole, this thesis indicates that measuring multiple aspects of vocabulary knowledge both receptively and productively may provide a much more accurate assessment of the relative efficacy of vocabulary learning tasks. Moreover, it suggests that different tasks may have a different effect on vocabulary knowledge.

ACKNOWLEDGEMENTS

I wish to express my gratitude to my supervisor, Paul Nation, for his keen insights and direction throughout this project. His enthusiasm and support helped to make this task a positive one. I am also indebted to Jim Dickie for guiding me through the statistical analyses, and to Dr. Jonathan Newton for acting as a reader.

I would also like to thank Brian Quinn, Jack Brajcich, Ken Brown, Stuart Gayle, and Mark Flynn for allowing me into their classes at Kyushu University to complete my research.

I am grateful to Mari Inoue and Chisato Mihara for their help with the Japanese translations, and the rest of my colleagues at Koran Women's Junior College for their support during the writing of this thesis.

I would also like to express my deepest appreciation to my wife, Machiko, for her patience, understanding, support, and love.

Finally, I would like to dedicate this work to my parents, Peter and Joan, for their love and support during these times and always.

TABLE OF CONTENTS

| | |
|---|--------|
| Abstract | ii |
| Acknowledgements | iv |
| Table of Contents | v |
| List of Tables | x |
| List of Figures | xii |
| <u>CHAPTER 1: INTRODUCTION</u> | 1 |
| 1.0 Introduction | 1 |
| 1.1 Research investigating vocabulary learning tasks | 2 |
| 1.2 Vocabulary knowledge | 3 |
| 1.3 How has vocabulary knowledge been measured? | 6 |
| 1.4 Thesis format | 9 |
| <u>CHAPTER 2: INCIDENTAL LEARNING FROM READING AND LEARNING FROM WORD PAIRS</u> | 10 |
| 2.0 Introduction | 10 |
| 2.1 Learning from word pairs | 11 |
| 2.2 L1 incidental vocabulary learning | 12 |
| 2.3 L2 incidental vocabulary learning | 16 |
| 2.4 The effects of context | 22 |
| 2.5 Implementation of the research | 26 |
| 2.5.1 Research questions | 26 |
| 2.5.2 Pilot studies | 27 |
| 2.5.3 Subjects | 29 |
| 2.5.4 Design | 29 |
| 2.5.5 Procedure | 32 |
| 2.5.6 Stage 1 procedure | 33 |
| 2.5.7 Stage 2 procedure | 33 |
| 2.5.8 Stage 3 procedure | 34 |
| 2.5.9 Stage 4 procedure | 35 |
| 2.5.10 Materials | 35 |
| 2.5.11 The number of target words | 35 |
| 2.5.12 Target words | 35 |
| 2.5.13 Sentence contexts | 36 |
| 2.5.14 Order and rating of contexts | 37 |
| 2.5.15 Word Pairs | 38 |
| 2.5.16 Time | 39 |
| 2.5.17 Dependent measures | 39 |
| 2.5.18 Productive knowledge of orthographic form | 40 |
| 2.5.19 Receptive knowledge of orthographic form | 41 |
| 2.5.20 Receptive recall of meaning | 41 |
| 2.5.21 Productive knowledge of grammatical functions | 42 |
| 2.5.22 Productive knowledge of syntax | 42 |
| 2.5.23 Productive knowledge of association | 43 |

| | | |
|--------|---|----|
| 2.5.24 | Receptive knowledge of grammatical functions | 44 |
| 2.5.25 | Receptive knowledge of syntax | 44 |
| 2.5.26 | Receptive knowledge of association | 44 |
| 2.5.27 | Receptive recognition of meaning and form | 45 |
| 2.6 | Case studies | 46 |
| 2.7 | Results | 48 |
| 2.7.1 | Productive knowledge of orthography | 49 |
| 2.7.2 | Receptive knowledge of orthography | 51 |
| 2.7.3 | Receptive recall of meaning and form | 51 |
| 2.7.4 | Receptive knowledge of meaning and form | 52 |
| 2.7.5 | Productive knowledge of association | 53 |
| 2.7.6 | Receptive knowledge of association | 54 |
| 2.7.7 | Productive knowledge of syntax | 55 |
| 2.7.8 | Receptive knowledge of syntax | 56 |
| 2.7.9 | Productive knowledge of grammatical functions | 57 |
| 2.7.10 | Receptive knowledge of grammatical functions | 58 |
| 2.8 | Discussion | 59 |
| 2.9 | Summary of main findings | 66 |
| 2.10 | Limitations | 67 |
| 2.11 | Rationale for the next experiment | 68 |

CHAPTER 3: LEARNING FROM WORD PAIRS AND LEARNING FROM GLOSSED SENTENCES

| | | |
|--------|---|----|
| | | 70 |
| 3.0 | Introduction | 70 |
| 3.1 | Research investigating the effects of context on vocabulary learning | 71 |
| 3.2 | What may be gained from learning in context? | 75 |
| 3.3 | Experiment 1 | 76 |
| 3.4 | Implementation of the research | 80 |
| 3.4.1 | Research questions | 80 |
| 3.4.2 | Pilot studies | 81 |
| 3.4.3 | Subjects | 82 |
| 3.4.4 | Design | 82 |
| 3.4.5 | Procedure | 83 |
| 3.4.6 | Stage 1 procedure | 84 |
| 3.4.7 | Stage 2 procedure | 84 |
| 3.4.8 | Stage 3 procedure | 85 |
| 3.4.9 | Materials | 85 |
| 3.4.10 | Target words | 85 |
| 3.4.11 | Nonsense words | 86 |
| 3.4.12 | Word pairs | 86 |
| 3.4.13 | Sentence contexts | 87 |
| 3.4.14 | Controlling for interlexical and intralexical factors that may affect vocabulary learning | 88 |
| 3.4.15 | Dependent measures | 89 |
| 3.4.16 | Productive and receptive measures of orthographic form..... | 91 |
| 3.4.17 | Productive and receptive measures of meaning and form | 91 |
| 3.4.18 | Productive and receptive measures of association | 92 |

| | |
|--|------------|
| 3.4.19 Productive and receptive measures of syntax | 93 |
| 3.4.20 Productive and receptive measures of grammatical functions | 93 |
| 3.5 Case studies | 93 |
| 3.6 Results | 94 |
| 3.7 Discussion | 99 |
| 3.8 Summary of main findings | 103 |
| 3.9 Rationale for the next experiment | 104 |
| CHAPTER 4: RECEPTIVE AND PRODUCTIVE LEARNING | 106 |
| 4.0 Introduction | 106 |
| 4.1 What is receptive and productive vocabulary knowledge? | 108 |
| 4.2 Do learners gain receptive knowledge before productive knowledge? | 111 |
| 4.3 Receptive and productive vocabulary size | 112 |
| 4.4 Learning from word pairs research | 117 |
| 4.5 The present study | 123 |
| 4.6 Tests | 125 |
| 4.7 Implementation of the research | 128 |
| 4.7.1 Research questions | 128 |
| 4.7.2 Pilot studies | 129 |
| 4.8 Experiment 3 | 129 |
| 4.8.1 Subjects | 129 |
| 4.8.2 Design | 130 |
| 4.8.3 Procedure | 130 |
| 4.8.4 Materials | 132 |
| 4.8.5 Target words | 132 |
| 4.8.6 Time | 132 |
| 4.8.7 Dependent measures | 132 |
| 4.8.8 Case studies | 132 |
| 4.8.9 Results | 133 |
| 4.9 Experiment 4 | 137 |
| 4.9.1 Subjects | 137 |
| 4.9.2 Design | 137 |
| 4.9.3 Sentence contexts | 138 |
| 4.9.4 Case studies | 139 |
| 4.9.5 Results | 140 |
| 4.10 Experiment 5 | 142 |
| 4.10.1 Subjects | 143 |
| 4.10.2 Design | 143 |
| 4.10.3 Target words | 143 |
| 4.10.4 Sentence contexts | 144 |
| 4.10.5 Case studies | 144 |
| 4.10.6 Results | 145 |
| 4.11 Discussion | 147 |
| 4.12 Summary of main findings | 154 |
| 4.13 Rationale for the next experiment | 155 |

| | |
|---|-----|
| CHAPTER 5: INVESTIGATING THE EFFECTS OF RECEPTIVE AND PRODUCTIVE LEARNING FROM WORD PAIRS ON COMPREHENSION AND WRITING | 157 |
| 5.0 Introduction | 157 |
| 5.1 How does vocabulary knowledge affect comprehension? | 158 |
| 5.2 Factors that may affect comprehension | 161 |
| 5.3 How does vocabulary knowledge affect writing? | 165 |
| 5.4 Implementation of the research | 167 |
| 5.4.1 Research questions | 167 |
| 5.4.2 Pilot studies | 168 |
| 5.4.3 Subjects | 169 |
| 5.4.4 Design | 169 |
| 5.4.5 Procedure | 170 |
| 5.4.6 Materials | 171 |
| 5.4.7 Target words | 171 |
| 5.4.8 Word pairs | 172 |
| 5.4.9 Time | 172 |
| 5.4.10 Dependent measures | 172 |
| 5.5 Case studies | 175 |
| 5.6 Results | 176 |
| 5.7 Discussion | 179 |
| 5.8 Summary of main findings | 184 |
| CHAPTER 6: CONCLUSION | 185 |
| 6.0 Introduction | 185 |
| 6.1 Review of the findings | 185 |
| 6.2 Testing | 190 |
| 6.3 The Involvement Load Hypothesis | 193 |
| 6.4 Conclusion | 196 |
| REFERENCES | 198 |
| Appendix A: Experiment 1 incidental learning contexts | 213 |
| Appendix B: Experiment 1 word pairs | 223 |
| Appendix C: Experiment 1 vocabulary knowledge tests | 224 |
| Appendix D: Experiment 2 word pairs | 230 |
| Appendix E: Experiment 2 glossed sentences | 231 |
| Appendix F: Experiment 2 vocabulary knowledge tests | 232 |
| Appendix G: Experiment 3 receptive word pairs | 242 |
| Appendix H: Experiment 3 productive word pairs | 243 |
| Appendix I: Experiment 4 glossed sentences | 244 |
| Appendix J: Experiment 4 writing task | 245 |
| Appendix K: Experiment 5 treatment | 246 |
| Appendix L: Experiment 6 receptive word pairs | 248 |
| Appendix M: Experiment 6 productive word pairs | 249 |

| | |
|--|-----|
| Appendix N: Experiment 6 picture description test | 250 |
| Appendix O: Experiment 6 true/false comprehension test | 252 |

LIST OF TABLES

| | | |
|----------|--|----|
| Table 1 | Description of treatments in Experiment 1 | 30 |
| Table 2 | Description of the dependent measures that were administered in Experiment 1 | 32 |
| Table 3 | Context rating for each repetition | 38 |
| Table 4 | Means and standard deviations of learning conditions on dependent measures in Experiment 1 | 49 |
| Table 5 | LSD test of differences in productive knowledge of orthography among the six groups | 50 |
| Table 6 | LSD test of differences in receptive knowledge of orthography among the six groups | 51 |
| Table 7 | LSD test of differences in productive knowledge of meaning and form among the six groups | 52 |
| Table 8 | LSD test of differences in receptive knowledge of meaning and form among the six groups | 53 |
| Table 9 | LSD test of differences in productive knowledge of association among the six groups | 54 |
| Table 10 | LSD test of differences in receptive knowledge of association among the six groups | 55 |
| Table 11 | LSD test of differences in productive knowledge of syntax among the six groups | 56 |
| Table 12 | LSD test of differences in receptive knowledge of syntax among the six groups | 57 |
| Table 13 | LSD test of differences in productive knowledge of grammatical functions among the six groups | 58 |
| Table 14 | LSD test of differences in receptive knowledge of grammatical functions among the six groups | 59 |
| Table 15 | Correlation coefficients between the number of repetitions and dependent measures in Experiment 1 | 62 |
| Table 16 | Description and order of the dependent measures in Experiment 2 | 90 |
| Table 17 | Means and standard deviations of learning conditions on dependent measures in Experiment 2 | 95 |
| Table 18 | Summary of multivariate analysis of variance on vocabulary tests | 95 |

| | | |
|----------|---|-----|
| Table 19 | Means and standard deviations of learning conditions on dependent measures for words with and without high frequency synonyms | 97 |
| Table 20 | Summary of multivariate analysis of variance on vocabulary tests for word pairs, learning from glossed sentences, and for the combined tasks..... | 98 |
| Table 21 | Mean percentage gains with the non-synonyms on the 10 dependent measures for the learning from word pairs group..... | 100 |
| Table 22 | Definitions of receptive and productive knowledge | 110 |
| Table 23 | Means and standard deviations on dependent measures for the receptive and productive word pairs tasks | 134 |
| Table 24 | Summary of multivariate analysis of variance on vocabulary tests in Experiment 3 | 134 |
| Table 25 | Means and standard deviations of combined receptive and productive scores for each aspect of knowledge in Experiment 3 | 135 |
| Table 26 | Summary of multivariate analysis of variance on combined receptive and productive scores for each aspect of knowledge | 136 |
| Table 27 | Means and standard deviations of learning conditions on dependent measures in Experiment 4 | 140 |
| Table 28 | Summary of multivariate analysis of variance on vocabulary tests | 141 |
| Table 29 | Means and standard deviations of combined receptive and productive scores for each aspect of word knowledge in Experiment 4 | 141 |
| Table 30 | Summary of multivariate analysis of variance on combined receptive and productive scores for each aspect of knowledge | 141 |
| Table 31 | Means and standard deviations on dependent measures for the reading and writing tasks in Experiment 5 | 145 |
| Table 32 | Summary of multiple analysis of variance on vocabulary tests | 146 |
| Table 33 | Means and standard deviations of combined receptive and productive scores for each aspect of word knowledge in Experiment 5 | 147 |
| Table 34 | Summary of multiple analysis of variance on combined receptive and productive scores for each aspect of knowledge | 147 |
| Table 35 | The range between the mean scores on the productive measures of orthography, association, meaning, and grammar | 153 |
| Table 36 | Means and standard deviations of picture description, and comprehension tests for the two learning conditions | 177 |
| Table 37 | Summary of multivariate analysis of variance on picture description and comprehension tests | 177 |

Table 38 Means and standard deviations of vocabulary tests for the two
learning conditions in Experiment 6 178

Table 39 Summary of multivariate analysis of variance on vocabulary tests 178

LIST OF FIGURES

Figure 1 What is involved in knowing the word *disabled* 5

Figure 2 Tests which measure the different aspects of word knowledge 7

CHAPTER 1

1.0 Introduction

The purpose of this thesis is to examine the effects of vocabulary learning tasks on word knowledge to determine whether different tasks result in different types of vocabulary knowledge. For example, contextualized tasks may be better suited to developing knowledge of syntax, collocation, and grammatical functions than decontextualized tasks. In contrast, decontextualized tasks may be superior at promoting knowledge of meaning and form, and association. In order to determine the effects of vocabulary learning tasks, we need to look at what aspects of word knowledge are involved in knowing a word and how they might be measured.

What does it mean to know a word? While researchers have examined this question in detail, and have proposed different criteria to describe vocabulary knowledge (see for example, Aitchison, 1994; Laufer, 1997; McCarthy, 1990; Miller, 1999; Nation, 1990, 2001; Richards, 1976), there is agreement that knowing a word involves much more than simply knowledge of its form and meaning. However, tests of vocabulary learning have almost exclusively focused on measuring knowledge of meaning and form, while ignoring the other aspects of word knowledge (Read, 1987; Schmitt, 1994a). Although in many situations time constraints may make it impractical to test multiple aspects, it may often be necessary when evaluating the effectiveness of vocabulary learning tasks. For example, direct learning tasks may contribute to greater gains in knowledge of meaning and form than indirect learning tasks which may promote larger gains in knowledge of syntax, and grammatical functions. However, no studies that have compared direct and indirect tasks have tested multiple aspects of word knowledge. Moreover, almost no research investigating the effects of other learning tasks has measured multiple aspects of knowledge.

The remainder of this chapter will briefly look at how word knowledge has been defined and measured, and research that has investigated vocabulary learning tasks. Further review of research examining specific learning tasks will be presented in Chapters 2 through 5.

1.1 Research investigating vocabulary learning tasks

Vocabulary acquisition research has investigated the effectiveness of many different learning tasks. Most often, two or more tasks are compared to determine which task is more effective. For example, looking up unknown words in a dictionary was found to be superior to learning from context (Knight, 1994), learning from word pairs was more effective than learning from context (Prince, 1996; Seibert, 1930), productive learning from word pairs may be more useful than receptive learning from word pairs (Griffin and Harley, 1996), and learning from word pairs and learning from glossed sentences were superior to learning from glossed texts and elaborated texts (Laufer and Shmueli, 1997). This type of research is valuable because it indicates that some tasks may be more useful to teachers and students than others. However, the weakness of the research is that task efficacy has usually been based solely on whether subjects gain knowledge of meaning and form.

While knowledge of meaning and form must be considered the most important aspect of vocabulary knowledge, in some cases it would seem to be a poor choice to measure the relative efficacy of tasks. For example, it should be expected that tasks that provide learners with both meaning and form such as learning from word pairs, the keyword technique, and learning from glossed sentences may be effective in promoting semantic knowledge. However, it might also be expected that the presence of context in the learning from glossed sentences task may promote larger gains in knowledge of syntax, and grammar than the other two tasks. If research investigating those tasks does not measure gains in all three aspects of knowledge, the results may be misleading. A

similar argument can be made when comparing direct and indirect learning tasks. Direct tasks tend to focus learners on linking meaning and form while indirect tasks tend to focus learners on understanding or using language. Therefore, direct tasks may be more effective at promoting semantic knowledge while indirect tasks may contribute to fuller word knowledge. If research does not test multiple aspects of knowledge, the results might be biased in favor of the direct tasks. Moreover, in indirect tasks, learners often do not gain knowledge of meaning; however, they may gain knowledge of other aspects of knowledge such as orthography, and grammatical functions more quickly.

While earlier studies have given insight into which tasks are best suited to developing knowledge of meaning and form, it would also be useful to investigate how each task affects different aspects of knowledge. This would provide a more accurate assessment of task efficacy. Not only would this give a better indication of which tasks are most effective but it would also help to show their strengths and weaknesses, what each task contributes to vocabulary knowledge, as well as which tasks may complement each other to improve learning. Since very few studies have investigated the effects of vocabulary learning tasks on multiple aspects of word knowledge, the results would be valuable to researchers who may gain a better understanding of vocabulary acquisition, testing vocabulary knowledge, and the interrelationships between the aspects. Moreover, it may also be of use to teachers and students to develop insight into the advantages and disadvantages of vocabulary learning tasks, and help them to select the tasks that best suit their needs.

1.2 Vocabulary knowledge

Since Richards (1976) described several aspects of word knowledge in his influential paper *The role of vocabulary teaching*, many more applied linguists have discussed what is involved in knowing a word (see for example, Aitchison, 1994; Laufer, 1997; McCarthy, 1990; Miller, 1999; Nation, 1990, 2001; Schmitt, 1993, 1994b, 1998, 2000).

Richards (1976) listed word frequency, syntactic behavior, form and its derivations, association, limitations on use, semantic value, and conceptual meaning as types of word knowledge. Although Richards did not set out to create an all-inclusive description of vocabulary knowledge (Meara, 1996), others have characterized it as such, and have built upon it to create a more complete vocabulary knowledge framework. Perhaps the most comprehensive descriptions of word knowledge were those proposed by Nation (1990, 2001). Nation (1990) listed eight different types of knowledge that are necessary to know a word, but later revised it adding a ninth aspect 'word parts' (Nation, 2001). He defined the nine aspects of vocabulary knowledge as follows:

1. Knowledge of the spoken form of a word
2. Knowledge of the written form of a word
3. Knowledge of the parts in a word which have meaning
4. Knowledge of the link between a particular form and a meaning
5. Knowledge of the concepts a word may possess and the items it can refer to
6. Knowledge of the vocabulary that is associated with a word
7. Knowledge of a word's grammatical functions
8. Knowledge of a word's collocations
9. Knowledge of a word's register and frequency

Nation further broke down each aspect into receptive and productive knowledge.

Receptive knowledge most often refers to recognition of vocabulary when reading or listening while productive knowledge refers to use of words in speech and writing. In order to master a word, a person must gain both receptive and productive knowledge of all nine aspects. This is a large task for all but the most frequent words for language learners due to the amount of knowledge that needs to be gained. Research has shown that learners tend to gain knowledge in small increments, and that they may build upon their previous gains through repeated exposures (Nagy, Herman, and Anderson, 1985). Thus vocabulary knowledge is not an all-or-nothing relationship, but a gradual process in which different types of knowledge are learned until potentially all aspects of knowledge are known for an item (Schmitt, 1993, 1994b). Gaining full knowledge of a word requires considerable gains in all nine aspects of knowledge, and therefore, a large number of words, particularly the less frequent ones may only be partially learned. It also seems likely that some aspects of knowledge are acquired before

others. Schmitt and McCarthy (1997), and Schmitt (1998, 2000) suggest that knowledge of form, and meaning may be gained before some of the other aspects such as collocation and register.

Figure 1 What is involved in knowing the word *disabled*

| Aspects of word knowledge R=receptive knowledge, P=productive knowledge | | | Receptive and productive uses for each aspect of knowledge |
|--|---|---|--|
| Form | spoken | R | recognize its phonological form |
| | | P | pronounce it correctly |
| | written | R | recognize its orthographic form |
| | | P | spell it correctly |
| | word parts | R | recognize that both <i>dis</i> and <i>able</i> have meaning, and relate both parts to the meaning of <i>disable</i> |
| | | P | select the correct word part in the suitable form in order to say or write it |
| Meaning | form and meaning | R | understand that <i>disable</i> represents a particular meaning |
| | | P | say or write <i>disable</i> in order to convey its meaning |
| | concept and referents | R | recognize the general concept of <i>disable</i> and know that it can have different meanings in different contexts |
| | | P | use <i>disable</i> in a variety of contexts to express its different meanings |
| | associations | R | recognize the words which are related to <i>disable</i> such as synonyms, antonyms, superordinates, and subordinates |
| | | P | say or write the words which are related to <i>disable</i> |
| Use | grammatical functions | R | recognize the grammatical patterns in which <i>disable</i> occurs |
| | | P | say or write <i>disable</i> correctly in context |
| | collocations | R | recognize the words in context which <i>disable</i> usually occurs with |
| | | P | say or write the words that often occur with <i>disable</i> |
| | constraints on use (register and frequency) | R | recognize how often <i>disable</i> is used and in what level of formality it occurs |
| | | P | use <i>disable</i> to express the level of formality in which it is intended |

(adapted from Nation, 2001)

Figure 1 shows just how much is involved in knowing a word. By breaking down each type of knowledge into its receptive and productive uses, Nation (2001) identified 18 components of word knowledge which could be measured to indicate how well words are known. The following section will examine how the different aspects of word knowledge have been tested.

1.3 How has vocabulary knowledge been measured?

Traditional vocabulary tests tend to measure knowledge of meaning and form, and ignore the other aspects of knowledge (Read, 1987; Schmitt, 1994a). In fact, many studies have relied on a single receptive or productive test of meaning and form as evidence that target words are known. Since knowledge of meaning and form is only one of the nine aspects in Nation's (2001) vocabulary knowledge framework, and both receptive and productive tests may be needed to determine full knowledge of a single aspect, it is likely that a very large number of studies have measured only partial knowledge of a single aspect.

Full knowledge of a word has rarely been tested. One reason for this is that there are no practical tests which measure all aspects of word knowledge. Since vocabulary tests tend to measure either receptive or productive knowledge but not both, 18 tests may be needed to measure a single meaning sense. If all of the meaning senses of a word were to be tested, Meara (1996) speculates that as many as 40 to 50 may be required. Obviously the time needed to create, administer, and grade this many tests makes it a difficult, and for most purposes an impractical task (Meara, 1996). In one attempt to measure full knowledge of 11 words in interviews, Schmitt and McCarthy (1997) report that on average it took more than two hours per person.

While it may be impractical to test for full knowledge of words, it is relatively easy to measure multiple aspects of knowledge. However, surprisingly few studies have tested more than one aspect. This may be because it is also important to measure one aspect

of knowledge with multiple tests to account for strength of knowledge. Studies have shown that multiple tests of a single aspect can discriminate between learners with partial knowledge and those with greater knowledge (Joe, 1998; Nist and Olejnik, 1995; McKeown, Beck, Omanson, and Pople, 1985; Paul, Stallman, and O'Rourke, 1990).

Figure 2 Tests which measure the different aspects of word knowledge

| Aspects of word knowledge R=receptive knowledge, P=productive knowledge | | | Tests that can measure the receptive and productive uses of each aspect of word knowledge |
|---|---|---|---|
| Form | spoken | R | dictation, hear the test word and choose L1 translation |
| | | P | read aloud, cued oral recall |
| | written | R | pronounce written words |
| | | P | word or sentence dictation |
| | word parts | R | break the word into parts |
| | | P | what do you call ... |
| Meaning | form and meaning | R | translate words into L1 |
| | | P | translate words into L2 |
| | concept and referents | R | translate the word in context |
| | | P | choose the words to translate L1 word |
| | associations | R | choose the words that you associate with test word |
| | | P | add to a list of associated words |
| Use | grammatical functions | R | choose the correct use of the word |
| | | P | use the test word in a sentence |
| | collocations | R | choose the correct use of the word |
| | | P | produce collocates of the test word |
| | constraints on use (register and frequency) | R | choose the word which represents UK use |
| | | P | produce a more formal word for X |

(adapted from Nation, 2001)

The few studies that have tested multiple aspects of knowledge have used different test formats. Schmitt (1998, 1999) used interviews to test for knowledge of meaning, association, collocation, and grammatical functions (1998), and orthography, meaning, association, and grammar (1999). Schmitt and Meara (1997) administered a test with a word association format to measure knowledge of grammatical functions and association, and Read (1993) designed the Word Associates Test to measure knowledge of collocation, association, and meaning. The studies by Read (1993), Schmitt (1998, 1999), Schmitt and Meara (1997) indicated that although there may be difficulties testing multiple aspects of knowledge, it is possible. Schmitt (1994b) pointed out the importance of the test format when measuring vocabulary knowledge. He noted that the method of measuring word knowledge must be carefully chosen to make certain that it tests for the intended type of knowledge. This is an important point. Often in vocabulary acquisition research, tests require subjects to demonstrate more than one aspect of knowledge to score correctly. However, this makes it uncertain why subjects score incorrectly; was it because they lacked one aspect of knowledge, another, or both? When testing multiple aspects of vocabulary knowledge, it is necessary to use tests that isolate the intended knowledge type. Nation (2001) listed different test formats which can measure the receptive and productive uses of each aspect of word knowledge. These are shown in Figure 2.

Figure 2 shows that there are test formats that can measure the different aspects of word knowledge. While difficulties such as learning effects and subject fatigue might make it impractical to use 18 tests to measure full knowledge of a word, many situations may not require such extensive testing. Since vocabulary learning is a gradual process in which some aspects may be gained earlier than others, it may not always make sense to test for full knowledge. Instead, researchers should focus on the method of learning, determine which aspects of knowledge may be gained, and measure those aspects if it is feasible. One situation in which this would apply is after learners complete vocabulary learning tasks. Many learning tasks deal with only one meaning sense for each word with learners often meeting words in a limited number of repetitions. In such tasks, learners may gain knowledge of some aspects such as form,

meaning, and grammatical functions but are unlikely to gain knowledge of others such as constraints on use, and collocations. Therefore, fewer tests may be needed to determine the extent of learning.

1.4 Thesis format

This thesis does not follow the format traditionally used in most Applied Linguistics theses. Instead of having one comprehensive literature review following the introduction, each experiment has a separate literature review. This type of format has been used previously with Griffin (1992) employing a similar approach. It allows the reader to get a clearer indication of the aims of each experiment, as well as the direction of the thesis. Although all of the experiments have a similar design, they investigate a number of different learning tasks, and a number of different issues. However, each experiment builds upon the findings of the previous one.

In broad terms, each of the experiments is concerned with the effects of vocabulary learning tasks on vocabulary knowledge. Experiment 1 (Chapter 2) seeks to determine the effects of incidental learning through reading on word knowledge, and compares it with direct learning from word pairs. The effects of context and repetition on vocabulary learning are also addressed. Experiment 2 (Chapter 3) examines the effects of context on vocabulary knowledge. In this chapter, learning from glossed sentences is compared with learning from word pairs. Experiment 2 also examines the effects of synonymy and learning burden on vocabulary knowledge. Chapter 4 presents experiments 3 through 5 which deal with the effects of receptive and productive learning on word knowledge. Experiment 3 compares receptive and productive learning from word pairs, and experiments 4 and 5 compare receptive learning from glossed sentences and productive learning from writing. Chapter 5 looks at how receptive and productive learning from word pairs affects communication as well as vocabulary knowledge. In the final chapter (Chapter 6), a discussion of the findings in all six experiments is presented and followed by the conclusion.

CHAPTER 2

2.0 Introduction

There are two approaches to vocabulary acquisition, intentional or direct learning, and incidental or indirect learning. In intentional vocabulary learning, a deliberate attempt is made to learn words. In incidental vocabulary learning, words are acquired through exposure to the language while the learner focuses on the usage. Vocabulary acquisition research has shown that L2 learners gain more knowledge of meaning from an explicit learning task, such as learning from word pairs, than from an incidental learning task (Laufer and Shmueli, 1997; Prince, 1996; Seibert, 1930). However, L1 researchers tend to agree that incidental learning of vocabulary from reading is responsible for significantly more vocabulary growth (e.g. Nagy, and Herman, 1985; Nation, 1982; Sternberg, 1987). L2 research has shown that both incidental and intentional vocabulary learning can produce significant gains in knowledge of meaning but research tends to be more supportive of direct methods (Horst, Cobb, and Meara, 1998). While comparing the different methods can provide valuable insight into vocabulary learning, it may be more useful to look at what each method contributes to vocabulary learning. One of the more commonly used and effective explicit tasks is word pairs, learning an unknown L2 word together with its L1 translation or L2 synonym. Perhaps the most common incidental task is learning vocabulary through reading.

L1 and L2 incidental learning studies have found that the number of times an unknown word is met in context affects whether its meaning will be acquired (e.g. Horst, Cobb, and Meara, 1998; Jenkins, Stein, and Wysocki, 1984; Saragi, Nation, and Meister, 1978). The majority of past research has equated gains in knowledge of meaning with acquisition for both direct and indirect tasks. Other aspects of knowledge which may also be acquired such as syntax, grammatical functions, and orthographic knowledge, have largely been ignored. The present study was designed to examine (a) whether

knowledge of syntax, grammatical functions, orthography, association, and meaning and form are gained through incidental learning of vocabulary from reading, and learning from word pairs, (b) whether Japanese intermediate learners gain significant vocabulary knowledge from reading for comprehension, (c) how repetition affects incidental learning, and how many repetitions are needed to acquire the different aspects of knowledge, and (d) how context affects the acquisition of word knowledge.

2.1 Learning from word pairs

Learning from word pairs is a traditional method of vocabulary learning that consists of memorizing foreign vocabulary items together with their synonyms or translations in a list. In the last thirty years, rote learning methods such as learning from word pairs have been out of favor with vocabulary acquisition researchers (Hulstijn, 2001; Tinkham 1989). Researchers tend to agree that the vast majority of L1 vocabulary is learned indirectly through reading or listening, and that relatively few words are learned from direct teaching (Nagy and Herman, 1985). This has led some to suggest that an emphasis should be placed on learning from context, and that intentional learning tasks should be discouraged (Crow, 1986; Krashen, 1989). However, Nation (1980, 1982, 2001) argues that learning from word pairs can be a fast and efficient method of acquiring L2 vocabulary, and reports that past research demonstrates that learning from word pairs is an effective way of learning L2 vocabulary.

Past research has shown that learners can acquire from nine to 58 L2 vocabulary items per hour with an average of 34 (Thorndike, 1908) to as many as 166 L2 words per hour from word pairs (Webb, 1962). Nation (1980) suggests that learners may be able to average from 30 to over 100 L2-L1 word pairs per hour without too much effort. In comparison to learning from context, learning from word pairs has been found to be a much more effective method of acquiring vocabulary (Prince, 1996; Seibert, 1930). While learners may have very different learning capacities to learn from word pairs (see Thorndike, 1908; Webb, 1962) there may also be differences between cultures. Tinkham (1989) found that Japanese students performed better when learning from rote

than American students, and suggests that teachers should take advantage of the fact that many students are very good at rote learning. Griffin and Harley (1996) found that learning productively from L1-L2 word pairs is more effective than receptively from L2-L1 word pairs, and Nation (1980) reports that L2-L1 or L1-L2 word pairs are more effective than L2 words paired with L2 synonyms.

In summary the studies investigating learning from word pairs demonstrate that it is an effective method of quickly acquiring the meaning and form of L2 vocabulary.

However, one limitation of the research is that the only aspect of knowledge that was measured was meaning and form. In all of the word pairs studies, vocabulary acquisition is equated with gains of meaning and form. Whether any of the other aspects of knowledge (Nation 1990, 2001; Richards 1976) have been acquired remains to be found. It should be expected that learning from word pairs would be a more effective method of acquiring meaning and form than incidental learning from reading. Learning from word pairs focuses learner's attention on the meaning-form link, whereas incidental vocabulary learning may focus learners on other aspects of knowledge such as syntax, associations, and grammatical functions as well as meaning and form. Two aims of the present study are (a) to find which aspects of knowledge are acquired when people learn from word pairs, and (b) to shed more light on which vocabulary learning task, learning from word pairs or incidental learning from reading, is more effective.

2.2 L1 incidental vocabulary learning

Research has demonstrated that L1 learners can incidentally acquire vocabulary from reading (see for example, Jenkins, Stein, and Wysocki, 1984; Nagy, Anderson, and Herman, 1987; Shu, Anderson, and Zhang, 1995). One of the most influential studies was done by Nagy, Herman, and Anderson (1985). In this study, eighth-grade students with above-average or average reading ability read one of two authentic passages taken from school books. One text was an exposition and the other was a narrative. A checklist test was administered three days before the treatment to determine whether

subjects had prior knowledge of the target words. Subjects completed two measures of vocabulary learning, an interview and a multiple-choice test after the reading. Both the interview and the multiple-choice test were designed to measure knowledge of meaning at three different levels of difficulty, minimal partial knowledge, partial knowledge, and complete knowledge. On the multiple-choice test the levels of difficulty were based on the closeness in meaning between the target word and the distractors. At the highest level of difficulty, the distractors represented concepts which were semantically similar to the target words. At the intermediate level, the distractors were of the same part of speech but semantically dissimilar to the target words. At the lowest level of difficulty, the distractors were not related to the target words semantically, nor were they of the same part of speech. In the interview test, interviewers rated the subjects' knowledge of each target word on a scale from zero to three. A score of zero was given when the subject was unable to demonstrate any knowledge of the target word. A score of one was given for minimal partial knowledge. Two points were awarded for substantial yet incomplete knowledge, and three points for full adult-like knowledge of the meaning of the target word.

The results demonstrated that incidental learning from authentic texts does occur for young L1 learners. Gains in knowledge of meaning were found to be significant at each level of difficulty for both the multiple-choice test and the interview. The probability of gaining full knowledge of the meaning of an unknown word from one exposure in context ranged from .10 to .15 depending on the type of text, and the type of test. Gains in partial knowledge ranged from .14 to .25, and gains in minimal-partial knowledge ranged from .19 to .21.

The gains are relatively small in comparison to other vocabulary learning methods. However, Nagy, Herman, and Anderson (1985) argue that when the amount of reading done by young learners is considered, incidental vocabulary learning can result in significant vocabulary growth. They estimate that young learners read approximately one million words in a year. Therefore, based on the results of the multiple-choice test that measured full knowledge of meaning, they suggest annual vocabulary growth from

incidental learning from reading ranges from 1,500 to 8,250 words for young L1 learners.

Nagy, Herman, and Anderson's (1985) study was valuable for four reasons. First, it provided solid evidence that incidental vocabulary learning does take place for young L1 learners. Second, it demonstrated that incidental vocabulary learning through reading should play an important part of L1 language learning. Third, it showed that different tests measure different degrees of knowledge. This is particularly useful as it provides an explanation of why some studies have been unable to demonstrate vocabulary learning; often tests are measuring whether subjects have gained full knowledge of words when they are more likely to have gained partial knowledge (Nagy and Herman, 1987). In those cases, learning is not found to occur even though it has taken place. Future research needs to employ multiple tests to measure the different degrees of learning, or at least use measures that are more sensitive to partial learning. Fourth, they developed a rationale for the huge vocabulary growth experienced by young L1 learners. Their 'incremental nature of word learning' theory explains how incidental vocabulary learning is likely to occur, as well as why many studies have been unable to find evidence of indirect learning from context (Nagy, Herman, and Anderson, 1985; Nagy, Anderson, and Herman, 1987). The theory claims the following points:

1. Vocabulary learning is a gradual process in which gains are made in small increments.
2. Repeated encounters with a word are likely to lead to full knowledge of that word.
3. Most contexts do not provide much information about the meaning of words.
4. Learners are unlikely to learn a word's meaning from one encounter in context. However, partial learning of the meaning of a word is likely to be gained.
5. Only one sense of a word's meaning is likely to be learned from meeting it in context once.
6. Tests which measure full adult-like knowledge of a word are unlikely to find evidence of incidental vocabulary learning.
7. Tests which are sensitive to partial gains in knowledge are likely to find evidence of partial gains in knowledge.

8. Substantial exposure to written language should lead to a significant amount of incidental vocabulary learning.

The results of another L1 study which examined the role repetition plays in learning words from context support Nagy, Herman, and Anderson's (1985) incidental vocabulary learning hypothesis. Jenkins, Stein, and Wysocki (1984) found significant incidental learning of word meanings from reading. Fifth-grade subjects read short paragraphs that had been written to convey knowledge of the meanings of low-frequency target words. Subjects met the target words two, six, or ten times over a period of ten days. However, two days before the treatment, subjects underwent 'familiarization training' which involved reading the target words in class. The meanings of the target words were not mentioned during that time. Four tests designed to measure full knowledge of meaning were completed two days after the treatment. The results suggest that the number of times words are met in context increases the probability of learning word meanings. On one test in which subjects had to write the meaning of an underlined target word that was shown in a sentence, the subjects who saw the target words in two contexts, six contexts, and ten contexts, outperformed the zero context control group by 16.7%, 29.3%, and 33.1%, respectively. Similarly, on a multiple-choice test in which subjects had to select the correct definitions of target words, the two context, six context, and ten context groups outperformed the control group by 14.2%, 26.2%, and 32.9%, respectively. On both tests, significant differences were found between the zero repetition control group and the six, and ten repetition groups. However, two repetitions did not significantly differ from zero repetitions on both tests. On a multiple-choice sentence completion test, and a reading comprehension test a similar trend was found with scores increasing as the number of repetitions increased. However, significant differences were found between only the zero and ten repetition groups. In an analysis of their data, Nagy, Herman, and Anderson (1985) report that the probability of being able to write the definition of a word after meeting it in context once was 5%. While the probability of selecting the correct definition was 10% from one exposure.

The findings helped to confirm that L1 learners do acquire vocabulary from reading, and also demonstrate that knowledge of meaning increases as the number of repetitions increases. The failure to find a significant difference between the two repetition and zero repetition control groups was likely to have been a result of the experimental design. Subjects in the two repetition group had an interval of nine days between context presentations, whereas subjects in the 10 repetition group were presented with contexts everyday. For the two repetition group, it is likely that gains in knowledge of meaning made in the first repetition were reduced or even nullified before the final repetition as backsliding or forgetting was likely to occur. Although spaced repetitions have been found to be more effective than massed presentations (Dempster, 1987), Landauer and Bjork (1978) report that long intervals between repetitions decrease the likelihood that learning will occur. A second weakness of the experimental design was that all four tests measured only full knowledge of meaning. Each group may have made gains in partial knowledge that was not found by the tests.

One limitation of the above L1 studies is that they equate gains in knowledge of meaning with vocabulary acquisition. Other aspects of knowledge such as grammatical functions, and syntax are also likely to be acquired from reading. One of the aims of the present study is to find whether knowledge of orthography, syntax, associations, and grammatical functions are acquired together with meaning and form from reading. A second goal is to investigate how repetition influences acquisition of the various aspects of knowledge.

2.3 L2 incidental vocabulary learning

Several studies have also found statistically significant incidental vocabulary learning gains from reading with ESL and EFL learners (Day, Omura and Hiramatsu, 1991; Dupuy and Krashen, 1993; Hulstijn, 1992; Pitts, White, and Krashen, 1989). Pitts, White, and Krashen (1989) found that ESL subjects who met target words from one to 27 times in a one hour reading for comprehension task had a mean gain of six percent of unknown words. This occurred despite the fact that over 50% of the subjects were

unable to complete the reading in the allotted time, and that the tests were not sensitive to partial knowledge. In two similar studies, Hulstijn (1992) reported an incidental vocabulary gain of eight percent for ESL learners, and Day, Omura, and Hiramatsu (1991: 543) report gains of six percent for Japanese high school students and 18 percent for Japanese university students who met target words at 'ample' frequencies in a reading task. Dupuy and Kashen (1993) found larger gains in word knowledge in an EFL study that combined watching a video and reading a text. Eight of the 30 target words appeared in both the video and the passage, and target words appeared between one and six times in the text. A gain of six words (20%) was found when testing for fuller knowledge of meaning, and seven words (22%) when measuring for partial knowledge. However, 30 percent of the subjects did not finish reading the text in the 40 minutes allotted indicating that there was the potential for higher vocabulary gains. Together these studies suggest that incidental vocabulary learning from reading does occur with ESL and EFL learners. However, none of the studies measured the effects of repetition, nor did they attempt to explain how learning occurred.

The first study to provide an explanation for L2 incidental vocabulary learning was done by Saragi, Nation, and Meister (1978). They suggest that as the number of times an unknown word is met in context increases, so does the probability of vocabulary acquisition. In their study, adult native speakers of English read Anthony Burgess's *A Clockwork Orange*. The novel contains 241 Russian slang words called nadsat. The nadsat words had a range of repetitions from 1 to 209. None of the subjects spoke Russian, and the glossaries containing English translations of the nadsat words were removed from the books before they were given to the subjects. The subjects were told that after they had finished reading the book they would be given a test of literary criticism and comprehension about the novel. Instead, a surprise multiple-choice test was given on 90 nadsat words. The test measured receptive knowledge of meaning. All 49 of the nadsat words that had 18 or more repetitions were included in the test. The remaining 41 items on the test had a range of repetitions from 1 to 17. Therefore, the test did not represent the natural distribution of the target words, and was biased towards words with a large number of repetitions. The average vocabulary gain was

68.4 words (75%) indicating that learners were able to incidentally acquire L2 vocabulary from reading. The correlation between the number of times each word occurred in the book and the relative learning gains was found to be 0.34. The size of the correlation confirms that repetition affects learning. However, it also indicates that other factors are involved. They suggest that the meaningfulness of the context, and the degree of similarity between the form of the L1 and L2 may affect acquisition. They go on to report that learners need to meet unknown words ten or more times in a reader for significant gains to occur. However, learning gains varied considerably from word to word. For example, 70 percent of the learners demonstrated knowledge for a word which occurred only once, 15 percent for a word which occurred nine times, and 40 percent for a word which occurred 42 times. While gains were more consistent at 10 or more repetitions, there was not a minimum number of repetitions which guaranteed that learning would occur. One target word with 96 repetitions was only learned by 40 percent of the subjects.

One criticism of the study is that it does not accurately portray second language learning (Horst, Cobb, and Meara, 1998). Horst, Cobb, and Meara argue that it would be more difficult for L2 learners to infer words from L2 contexts than for native speakers to infer L2 words from an L1 context. Their point is a good one. The adult native speakers in the study are likely to have had fuller knowledge of the running words in the text than most L2 learners would have when reading. This would have given them an advantage over L2 learners trying to guess unknown words, and may account for the much larger gains made in this study. A second reason for the large gains may be that the adult native speaker subjects were more skilled in guessing a word from context. The ability to guess a word from context can greatly affect vocabulary acquisition (Jenkins, Matlock, and Slocumb, 1989).

Despite this weakness, the Saragi, Nation, and Meister study may provide a reasonable assessment of L2 incidental vocabulary learning under ideal conditions. The ratio of known running words to unknown nasdat words in *A Clockwork Orange* is 94 to 6. This represents the minimal conditions which most L2 vocabulary researchers believe

is necessary for L2 incidental vocabulary learning to take place (see for example, Laufer, 1992; Liu Na and Nation, 1985; Nation, 1990; Schmitt, 2000). Liu Na and Nation (1985) report that L2 learners need to know a minimum of 95 percent of the words in a text to be able to infer unknown words. Extensive reading programs using graded readers may fulfill these conditions (Nation and Wang, 1999). However, a more recent study examining the degree of coverage required to gain knowledge of unknown words suggests that 98 percent coverage may be needed to comprehend a text without assistance (Hu and Nation, 2000).

Horst, Cobb, and Meara (1998) replicated Saragi, Nation, and Meister's (1978) *A Clockwork Orange* methodology with L2 learners. Subjects followed along in a graded reader as it was read. The reading was completed in six one-hour classroom sessions completed over a period of ten days. Subjects were tested for their knowledge of 45 low frequency words in a pre-test post-test design. Two tests were used to measure the subjects' knowledge of the target words, a multiple-choice test which measured receptive knowledge of meaning, and a word association test in which subjects had to choose the two words which had a semantic link from a set of three words. Vocabulary gains of 22 percent, and 16 percent were made on the multiple-choice test, and on the association test, respectively. The results are more impressive considering that it is likely that the text that was used was probably too difficult for many learners. Based on the results of the Vocabulary Levels Test (Nation, 1990) at the 2000 word level, they estimate that the subjects knew an average of 1203 of the most frequent 2000 words. This suggests that the 2000 word level reader used in the study contained a large number of unknown words. The researchers attributed the gains to the frequency of occurrences in the reader. A correlation of 0.49 was found between the number of times each word occurred in the book and the gains made by the students. Furthermore, they suggest that large learning gains are likely to occur for words which were repeated eight or more times. Acquisition of words which were repeated fewer than eight times was much less predictable with some cases of negative gains found between the pre and post-tests. The size of the gains varied considerably from word to word. The size of the correlation also suggests that other factors are involved. Two factors which they

believe may have influenced learning were pictures in the book, and the part of speech of the target words. Many concrete nouns were found to have high gain scores, and the target words which were shown together with pictures in the book also had very high gain scores despite only occurring five times in the text.

Horst, Cobb and Meara (1998) suggest that Nagy, Herman, and Anderson's (1985) incidental vocabulary learning hypothesis is not applicable to beginning ESL learners due to the much smaller volume of reading which occurs. Their argument is that one million words are roughly equal to 50 graded readers. At the rate of five words learned per reader, the annual gain is a very disappointing 250 words. However, in this argument they ignore the following important points:

1. The gain of 5 words was taken from a sample of 23 unknown words, a 22 percent gain. Their argument holds true if no other words in the 21,132 word text are unknown which is extremely unlikely considering the mean score of 1200 on the Vocabulary Levels Test.
2. Their argument also ignores partial learning gains which were not large enough to be measured as known on the post-tests that they used.
3. Gains may still have been made for the low frequency words which learners demonstrated knowledge of on the multiple-choice pre-test. It is unlikely that the learners had full knowledge of all of those words. A second test that measured fuller knowledge of meaning and form may have been able to show this.
4. Gains in word knowledge will accrue from text to text. Unknown words which were met in context once or twice and only partially learned may be read again in another text. If there are enough repetitions between texts the words should eventually be learned.

Overall, the results support L2 incidental vocabulary acquisition findings. The gains found in this study were larger than in some of the previous *A Clockwork Orange* replication studies (Dupuy and Krashen, 1993; Pitts, White, and Krashen, 1989) despite

using a more strictly controlled design. Furthermore, it supports Saragi, Nation, and Meister's (1978) findings that repetition plays a significant role in L2 incidental vocabulary learning from reading.

Two studies that measured the effect of repetition on vocabulary learning and used multiple tests at different sensitivities to assess knowledge of meaning had contrasting results. Rott (1999) examined how two, four, and six repetitions affected incidental gains in knowledge of meaning. Intermediate ESL students read a number of contrived passages and were then tested immediately after the treatment. Two measures at different sensitivities were used to assess receptive knowledge of meaning. The results suggest that subjects who encountered words twice produced significant learning gains (43%, 19%). Six repetitions (61%, 45%) produced significantly more learning gains than two (43%, 19%) or four (44%, 28%) repetitions. There was little difference between two and four repetitions. Rott suggests that six repetitions may be enough for considerable lexical gains to occur. However, it should be noted that the texts in the study were created with enough contextual clues for the subjects to be able to infer the meaning of the target words. In an authentic text this is not always the case (Beck, McKeown, and McCaslin, 1983). Therefore, the number of repetitions needed may be higher than the six Rott proposes.

Hulstijn, Hollander, and Greidanus (1996) used two receptive recall tests to assess gains in knowledge. In the first test the subjects had to write the target words' meanings, and in the second they met the target words in context and then had to write the meaning of each target word. The results showed that the subjects were able to recall more meanings for the words that they encountered three times; however no significant differences between target words met once or three times were found. They attribute the lack of a repetition effect to the learners ignoring the unknown target words or making incorrect inferences. It should be noted though that the text was authentic, and the researchers determined through pilot testing that it was 'extremely difficult to infer the exact meaning' of the target words from the context (p. 330).

Taken as a whole the L2 studies provide considerable evidence that repetition affects incidental vocabulary learning. However, two points remain unclear. First, how many repetitions are needed for acquisition to occur? Saragi, Nation, and Meister (1978) suggest 10, Horst, Cobb, and Meara (1998) suggest 8, and Rott (1999) reports that 6 may be enough. Second, to what degree are the other aspects of knowledge such as syntax, grammatical functions, association, and orthographic knowledge learned? The only study to measure a knowledge type besides meaning was Horst, Cobb, and Meara (1998) who also found gains in knowledge of association. Schmitt (1998) may have completed the only study which has examined this question but was unable to find any evidence to support an implicational scale of knowledge types.

2.4 The effects of context

It seems likely that the contexts may have been one of the reasons for the contrasting results in Rott (1999) and Hulstijn, Hollander, and Greidanus (1996). This may be the case in many incidental learning studies because a number of different types of contexts have been used. For example, Herman, Anderson, Nagy, and Pearson (1987) use four different types of context in a study of incidental vocabulary learning. They use a 1,230 word passage from a text and three edited versions with varying degrees of implicit and explicit clues about the target words. Other contexts used in incidental learning of vocabulary experiments were a novel (Saragi, Nation, and Meister, 1978), a graded reader (Horst, Cobb, and Meara, 1998), specially constructed paragraphs (Jenkins, Stein and Wysocki, 1984), and narrative and expository texts of about 1000 words (Nagy, Herman, and Anderson, 1985). Studies on learning from context have used single sentences (Dempster, 1987; Laufer and Shmueli, 1997), multiple sentences (Dempster, 1987), L2 sentences together with their L1 translations (Griffin, 1992), L1 sentences containing the L2 target vocabulary (Pickering, 1982), three sentences one of which was a definition (Gipe and Arnold, 1979), and L1 and L2 glossed passages (Laufer and Shmueli, 1997). The wide range of contexts may lead to misinterpretation of results. Would students learning vocabulary from a specially constructed passage containing many clues about a target word's meaning learn as easily from an isolated

sentence? Probably not (Beck, McKeown, and McCaslin, 1983; Herman, Anderson, Nagy, and Pearson, 1987). Nation (1982) points out the need for a clear definition of the term context. In the present study each individual context was rated on the amount of information it provided about a target word. By rating each context on the information that is provided about target words, the effects of context can be isolated. Context may be one of the reasons why some words have been acquired after very few meetings and others have needed many more repetitions to be acquired. One of the aims of this study is to find how context affects the acquisition of different aspects of knowledge. A comparison of two groups, one meeting words in more informative contexts, may shed light on this.

There are four major differences between the present study and past incidental vocabulary learning studies. First, each context in this study was rated on the amount of information it provided about a target word's meaning. Rated contexts may give a more accurate assessment of vocabulary learning because they take into account that some contexts provide little, or misleading information about a word's meaning while others provide many clues. One weakness of previous L2 research is that little or no information has been provided about the contexts. This is perhaps due to the length of the texts, and the large number of occurrences of target words in those studies.

Another difference between the present study and prior studies is that more aspects of knowledge were measured in the present study. In most vocabulary acquisition research, acquisition has been defined by whether subjects were able to demonstrate knowledge of a word's meaning. Other aspects of word knowledge which may also be gained from meeting an unknown word when reading such as syntagmatic associations, and grammatical functions, have until now been ignored. Determining whether multiple types of word knowledge are acquired from reading would be useful for two reasons. First, it could help teachers to better understand the interrelationships between the knowledge types. This could help teachers to improve their vocabulary learning programs. For example, if it were found that when intermediate learners acquired the meaning of an item from word pairs they also gained partial knowledge of another

aspect such as association but did not gain any knowledge of syntax, then teachers could combine learning from word pairs with vocabulary learning activities that develop syntactic knowledge. Second, measuring multiple aspects may help researchers to better understand the vocabulary acquisition process. Schmitt (1998, 2000), and Schmitt and McCarthy (1997), have suggested that there may be a developmental hierarchy of knowledge types. Although Schmitt (1998) was unable to find evidence to support this, his argument about the interrelationships between word knowledge types is compelling. He suggests that knowledge of form and one meaning sense are likely to be acquired first, and the last aspects of knowledge to be acquired may be collocation and stylistic knowledge.

A third difference is that in the present study, two measures at different sensitivities were used to assess each aspect of knowledge. Previous studies have tended to test for full knowledge of an aspect with a single measure. Using multiple tests to measure one knowledge type increases the possibility that at least partial learning may be found (Nagy, Herman, and Anderson, 1985).

A fourth difference between this study and past research is that most studies investigating incidental vocabulary acquisition have used low frequency words as the target items. However, using nonsense words has several advantages over low frequency words. First, it ensures that subjects have no prior knowledge of the target words. Finding target words that no subjects have met previously is very difficult. Often there are target words that some students know. To avoid this, very infrequent target words are used. Unfortunately, infrequent words do not often occur in text with very frequent words. Therefore, it is very difficult for researchers to use authentic text and low frequency words. Another advantage of using nonsense words is that there is no need to use a pre-test to measure for prior knowledge of the target words. Studies that use low frequency words use pre-tests to control for prior knowledge. If it is found that some subjects have prior knowledge of a target word, then the item can be omitted from the study or it can be taken into account when formulating the results. There are several weaknesses with this approach. Some pre-tests may not be sensitive enough to

measure partial knowledge. Learners may have met a word in context but only have a vague sense of its meaning. They may have full or partial grammatical or orthographic knowledge though. There may also be a small learning effect from doing a pre-test. Certainly, subjects would gain knowledge of form from seeing target words in a pre-test. Another problem with the use of pre-tests is that subjects may be tipped off that the focus of the study is vocabulary acquisition. The present study investigated incidental vocabulary acquisition. If learners were aware that they would be tested on their vocabulary knowledge, it would have been very unlikely that any vocabulary gains that took place were incidental.

One disadvantage of using nonsense words may be that because they are not authentic vocabulary items, they do not represent authentic vocabulary learning. However, all of the nonsense words created in this study resembled authentic vocabulary items orthographically and phonetically. In pilot studies, learners stated that they believed the items to be authentic. Since subjects in the experiment were not aware that nonsense words were used, authentic vocabulary learning should have occurred.

Another criticism may be that by substituting nonsense words for real words there could be a carry over learning effect once learners were aware of the meaning of a nonsense word. This, however, could be true anytime a learner learns a synonym for a word. For example, in the experiment the nonsense word 'ancon' is substituted for 'hospital'. Would the learning of 'ancon' be any different from a subject learning 'hospice' or 'sanatorium' words which have similar meanings to hospital? If 'hospice' or 'sanatorium' were used instead of 'ancon', we could not be certain that the subjects had no knowledge of the target word making nonsense words such as 'ancon' a better choice.

These three differences, rated contexts, multiple tests at different sensitivities, and using nonsense words as the target words should help to clear up some of the uncertainties in the prior studies as well as increase our understanding of the vocabulary acquisition process.

2.5 Implementation of the research

2.5.1 Research Questions

The present study will address the following questions:

1. How many exposures to a word met in context are needed for learners to gain knowledge of the following:
 1. productive knowledge of form
 2. receptive recall of meaning
 3. productive knowledge of grammatical functions
 4. productive knowledge of syntax
 5. productive knowledge of association
 6. receptive knowledge of grammatical functions
 7. receptive knowledge of syntax
 8. receptive knowledge of association
 9. receptive knowledge of meaning
 10. receptive knowledge of form

Research has shown that the more learners meet a word in context the more likely they are to gain knowledge of its meaning (e.g., Horst, Cobb, and Meara, 1998; Saragi, Nation, and Meister, 1978). It is likely that the other types of word knowledge would follow the same principle. Measuring word knowledge after learners have met target words n times may show how many repetitions are needed to gain knowledge for each type of word knowledge.

2. Which knowledge types are gained through learning from word pairs?

Studies examining learning from word pairs have only measured knowledge of meaning. Aspects of knowledge such as syntax, grammatical functions, association, and orthography have been ignored. However, it is likely that at least partial knowledge of those aspects would also be acquired.

3. Is incidental vocabulary learning as effective as learning from word pairs?

Laufer and Shmueli (1997), Prince (1996), and Seibert (1930) found that EFL students learning from word pairs gained significantly more knowledge of meaning than from learning vocabulary from reading. This could be expected because learning from word pairs is a meaning based task. Incidental vocabulary learning may produce greater knowledge of other knowledge types such as syntax and grammatical functions.

4. What effect does context have on knowledge of orthography, syntax, association, meaning and form, and grammatical functions?

Beck, McKeown, and McCaslin (1983) report that in some contexts vocabulary can be easily acquired while in others it is much more difficult. By rating each context on the information that is provided about target words, the effects of context can be isolated. Comparing vocabulary gains between subjects who meet target words in more informative contexts and less informative contexts may demonstrate that context has a significant effect on vocabulary acquisition.

5. Do Japanese foreign language learners acquire incidental vocabulary from reading as Day, Omura, and Hiramatsu (1991) suggest?

Day, Omura, and Hiramatsu (1991) found that Japanese EFL learners could incidentally acquire knowledge of meaning from reading. The present study may confirm their results using a more strictly controlled experimental design.

2.5.2 Pilot studies

Pilot studies were conducted with Japanese learners studying English as a foreign language to determine the following:

1. At which repetitions are aspects of word knowledge likely to be acquired?
2. Do the nonsense words create any interlanguage problems for Japanese learners such as phonetic similarities to an L1 or L2 item?
3. Are the learners able to understand the contexts to be used in the experiment?
4. Are the learners able to understand how to complete the various tests?
5. How much time do learners need to understand the contexts?
6. How much time do learners need to complete the treatment and tests?
7. How many learners are likely to be needed in each group for the findings to be significant?
8. Is there a learning effect from completing the tests?
9. Is there a fatigue factor involved when doing the tests?
10. Are the test items suitable?

As a result of the pilot studies, the following changes were made to the study.

1. The nonsense words were modified. Japanese learners have difficulty with the pronunciation and spelling of words which contain 'l' and 'r'. Many target words in the pilot studies contained those letters and thus contributed to very low scores on the productive knowledge of orthographic form test.
2. Several contexts used in the original pilot studies were omitted from the study because learners had difficulty understanding vocabulary in them.
3. To decrease the time taken to complete the tests and to aid in understanding how to complete the syntax and association tests instructions were written in Japanese as well as English for those tests.
4. Four minutes was found to be sufficient time to read one page of ten contexts.

5. It was found that learners were able to complete the treatment and tests in each group well within 90 minutes, the maximum amount of time possible.
6. Test items were modified to more accurately assess the type of knowledge to be tested, and to avoid providing clues for any of the following tests.

2.5.3 Subjects

The population sample consisted of 169 volunteers learning English as a foreign language in Fukuoka, Japan. The subjects were selected from four second-year EFL classes at Kyushu University. All of the subjects had studied English for a minimum of seven years, and had scored 80 percent or higher on the 2000 word level (a frequency band containing the 1001st to 2000th most frequent words) of Version 1 of the Vocabulary Levels Test (Schmitt, 2000). Their mean score was 27.5 indicating that they were well in control of that level, and that they had receptive knowledge of almost all of the 2000 most frequent words (Schmitt, Schmitt, and Clapham, 2001). 146 of the subjects were randomly assigned to the experimental and comparison groups. The control group consisted of the other 23 subjects. They were selected from one second-year EFL class at the same university.

2.5.4 Design

Five experimental groups, one comparison group, and one control group were used in this study to examine the effects of incidental learning through reading and learning from word pairs on vocabulary learning. Each experimental group completed a vocabulary comprehension task. The task involved reading a number of contexts containing ten target words (see Appendix A). The number of occurrences of each target word was different for each group. A surprise vocabulary test measuring different aspects of vocabulary knowledge was administered after the treatments were finished. The results of the groups were then compared to determine whether

significant learning could be attributed to the number of times the subjects met the target words in context. The comparison group completed a direct learning task, learning from word pairs (see Appendix B). No contexts were provided in the learning from word pairs treatment. They also took the ten tests to determine which aspects of knowledge were gained. Since the comparison group was given the meanings of the target words in the word pair task, how a gain of meaning affects the other aspects of knowledge could be examined. The control group (C) was added to the experiment after the treatments and tests had been completed. The control group did not complete any of the learning tasks and never saw any of the target words. The control group completed the productive knowledge of orthography test to determine whether gaining productive knowledge of form was the result of meeting the target words in the learning tasks or the result of the phonological prompt in the test.

Table 1 Description of treatments in Experiment 1

| Group | Repetitions / Task | Rated Contexts | Time |
|-------|--------------------|-------------------------------|------------|
| E1 | 1 | 1 | 4 minutes |
| E2 | 3 | 1, 2, 3 | 12 minutes |
| E3 | 3 | 1, 9, 10 | 12 minutes |
| E4 | 7 | 1, 2, 3, 4, 5, 6, 7 | 28 minutes |
| E5 | 10 | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 | 40 minutes |
| WP | word pairs | No contexts | 4 minutes |
| C | control | No contexts | 0 |

Experimental group 1 (E1) met each target word once, experimental group 2 (E2) saw each target word three times, experimental group 3 (E3) also met each target word three times, and experimental groups 4 (E4) and 5 (E5) saw the target words seven and ten times, respectively. The number of times that experimental groups 2, 3 and 4, and 5 viewed target words was determined in pilot studies. Experimental groups 2 and 3 both met target words three times but the contexts were different. E2 met target words in the three most informative contexts, whereas E3 met the target words in the most informative context, and the two least informative contexts. This was done to try to determine the effects context has on incidental learning. The comparison group (WP)

met the target words together with their L2 meanings for the same amount of time, four minutes, as E1.

Target words were read in short texts with an average of 14 words per text. The following example is for the target word 'denent' (remember): 'Perhaps I had a Christmas with my mother once, but I do not denent it.' Each text was taken from one of six graded readers. Pilot tests were used to confirm that the only unknown words in the texts were the target words. Limiting the number of unknown words to only the target words ensured that subjects would have sufficient time to focus on the target words. It is also likely to mirror incidental vocabulary learning from graded readers where learners should encounter only a small percentage of unknown words. Target words were presented to subjects in print. Each page presented ten contexts, each context containing a different target word. Therefore, after reading one page, subjects had seen each target word once.

Subjects were not given any instruction on incidental vocabulary learning strategies prior to the experiment. Neither were they told the nature of the study until after completion of the tests. The subjects were never told that they would be tested on vocabulary. This was to discourage the subjects in the experimental groups from concentrating solely on learning the unknown vocabulary. Since the present study investigated incidental vocabulary learning, the aim was for subjects to focus on reading comprehension.

After the treatment was finished, subjects were given the tests. Each test was on one page. Subjects were given as much time as they needed to finish a test. When a subject finished a test, it was collected and then the next test was handed out. Vocabulary learning was measured in ten ways. Table 2 shows the tests that were used and what they measured.

Table 2 Description of the dependent measures that were administered in Experiment 1

| Test order | Test type | Knowledge measured |
|------------|--------------------------------|---|
| 1 | Spelling | productive knowledge of orthographic form |
| 2 | Multiple-choice | receptive knowledge of orthographic form |
| 3 | Translation | receptive recall of meaning and form |
| 4 | Sentence construction | productive knowledge of grammatical functions |
| 5 | Write a syntagmatic associate | productive knowledge of syntax |
| 6 | Write a paradigmatic associate | productive knowledge of association |
| 7 | Multiple-choice | receptive knowledge of grammatical functions |
| 8 | Multiple-choice | receptive knowledge of syntax |
| 9 | Multiple-choice | receptive knowledge of association |
| 10 | Multiple-choice | receptive knowledge of meaning and form |

2.5.5 Procedure

The present study was conducted in four stages. Stages 1-3 were carried out over two weeks in second-year EFL classes at Kyushu University in Fukuoka, Japan. The fourth stage was conducted approximately four months after the earlier stages had been completed. In the first and third stages of the experiment, subjects were involved in learning English as a Foreign Language with the researcher as their instructor. During the first stage, students were informed that part of the following class would involve research being undertaken by their instructor. Consent forms were given to all potential subjects, and their contents and any questions were discussed. Completed consent forms were collected later that class from learners willing to participate in the experiment. In the second stage of the experiment, six subjects who did not take part in stage three were involved in case studies outside of regular class time. Each subject was assigned to one of the five different experimental conditions, and to the comparison condition (WP). They were asked to complete the third stage of the experiment. While they completed the treatment and dependent measures they were asked to think aloud to provide insight into subjects' thought processes during the experiment. Their results were not included in stage three. During the third stage of the experiment, the subjects completed the treatments and dependent measures, while

other students in the classes worked on another part of their coursework. In the fourth stage the control group completed the productive knowledge of orthography test. The experimental procedure was implemented over the following four stages:

- | | |
|---------|--|
| Stage 1 | The Vocabulary Levels Test was administered to determine the vocabulary size of potential subjects. Subjects were randomly assigned to treatment groups. |
| Stage 2 | The case studies were conducted. |
| Stage 3 | Subjects participated in the treatments, and the dependent measures were administered. |
| Stage 4 | The control group completed the productive knowledge of orthography test. |

2.5.6 Stage 1 procedure

The Vocabulary Levels Test (Schmitt, 2000) was administered to potential subjects. At the outset, the test was explained and any questions regarding the test were answered to make certain that the instructions were completely understood. Subsequently, volunteers who had scored 80 percent or higher at the 2000 word level were randomly assigned to the five experimental groups and the comparison group. The Vocabulary Levels Test was used to ensure that all subjects in the study had a minimum vocabulary size. The test and the pilot studies made it reasonably certain that the subjects would know all of the words presented in the texts to the experimental groups.

2.5.7 Stage 2 procedure

Case studies were done with six volunteers who scored 80 percent or higher at the 2000 word level on the Vocabulary Levels Test. None of the subjects were involved in any of the classes that participated in Stage 3. The purpose of the case studies was to gain insight into the subjects' thought processes when learning unknown words. The six subjects were randomly assigned to do one of the six different treatments. Each subject

was asked to think aloud as they did the treatment and the dependent measures. The procedure in stage 2 was identical to stage 3 except that the subjects verbalized their thoughts. After the dependent measures were completed, subjects were asked a number of questions to clarify any misunderstandings which had occurred. During stage 2 each subject was audio-recorded.

2.5.8 Stage 3 procedure

At the outset, instruction on how to complete the treatments was given. Their instructions were as follows:

You will be given a few pages to read. Some people will get more pages than others. You have four minutes to read each page. After four minutes that page will be taken away so you will not be able to read it again. You cannot use a dictionary. When you have read every page you will be given a few questions to find whether you could understand what was written.

Any questions that subjects had were addressed to make certain that the instructions were completely understood. The subjects were never told that they would be tested on vocabulary.

The pages to be read were placed face down on the subjects' desks. E1 participated in a treatment with one page of contexts, E2 and E3 subjects received three pages, E4 and E5 received seven and ten pages, respectively, and WP received one page with the target words and their L2 synonyms. Since the first test needed to be done by all subjects at the same time, groups began at different intervals so that all of the groups finished the treatment together. At the starting point, students were told to turn over the first page and to read it. After four minutes an alarm sounded, and students were told to turn over that page and begin reading the next page. The first page was then collected. This continued every four minutes until the treatment had been completed. At this time, the first dependent measure was administered. Subjects had as much time

as they needed to finish a test. When a subject had finished a test it was collected, and then the next dependent measure was administered.

2.5.9 Stage 4 procedure

During the fourth stage, the control group completed the productive knowledge of orthography test. The test was identical to the one used in stage 3.

2.5.10 Materials

2.5.11 The number of target words

Ten target words were used in this experiment. The number of target words was determined during pilot studies. Factors, which were considered when establishing the number of target words, were subject fatigue, time needed to complete the treatment and tests, and the number of target words needed to provide significant results.

2.5.12 Target words

Ten nonsense words were created and used as target words in the experiment. All of the nonsense words were two syllables. Seven of them were five letters long, and three of them were six letters long. The nonsense words were as follows: ancon, cader, dangy, denent, faddam, hodet, masco, pacon, sagod, and tasper.

Ten words were selected from six graded readers. The words were then deleted from the text and replaced with the nonsense words. Six of the words were nouns and four of the words were verbs. The ratio of six nouns to four verbs was used because nouns and verbs are the most common parts of speech found in natural text, and the 6:4 ratio approximates their proportional frequency of occurrence (Kucera and Francis, 1967). The ten words were the following: evening, face, hospital, lunch, remember, street, train, visit, wear, and write. Only one sense of a target word's meaning appeared in the

contexts. For example the meaning of 'face' in each occurrence meant 'the front of a person's head' not another meaning such as 'a person's expression' or 'the front of a cliff.' Similarly, the part of speech of each target word never changed from context to context. For example, although 'face' and 'lunch' occur as nouns and verbs they were always presented in the selected contexts as nouns.

2.5.13 Sentence contexts

Subjects in the experimental groups met target words in short contexts (see Appendix A). A context was either one or two sentences long and averaged approximately 14 words. All of the contexts used in this experiment were selected from graded readers. Selecting authentic sentences increased the ecological validity of the experiment. Contexts were taken from the following graded readers: *The Elephant Man*, *Lord Jim*, *Agatha Christie*, *Chemical Secret*, *The Garden Party and Other Stories*, and *Decline and Fall*. All of the graded readers were from the Oxford Bookworm series. Three factors were considered when selecting a context, the number of words, the frequency of each word in the sentence, and the ease in comprehending the sentence. Contexts were limited to those containing 43 or fewer words. There were three reasons for this. First, longer contexts increased the amount of time needed to complete the treatment. Second, the probability of finding sentences made up of words which all of the subjects knew was reduced as the number of words in the sentences increased. Third, shorter sentences would probably be more easily understood. This experiment examined the effect of repetition on vocabulary learning. If subjects were unable to comprehend a context, it would not be clear whether that occurrence contributed to a learners' vocabulary knowledge. The ease with which a target word could be inferred was not a factor when selecting contexts.

Extensive pilot testing was used to ascertain that subjects were familiar with all of the running words in the contexts. Contexts that contained words that were unknown to any subject in the pilot tests were not used in the study. Controlling the running words in the contexts ensured that the learners would quickly understand the sentences.

2.5.14 Order and rating of contexts

The contexts were rated by the experimenter and a second native speaker according to the following scale:

1. Extremely unlikely that the target word can be guessed correctly. The text contains no contextual clues and may be misleading.
2. It is unlikely that the exact meaning of the target word can be inferred. However, information in the context may lead to partial knowledge of the target word's meaning.
3. Information in the context may make it possible to infer the meaning of the target word. However, there are a number of choices. Subjects may gain partial knowledge.
4. Subjects have a good chance of inferring the meaning correctly. There are few meanings which are logical apart from the correct meaning. Subjects should gain at least partial knowledge.

The order of the contexts was determined by their ratings. The most informative context for each target word was the first repetition, the second most informative context was the second repetition while the least informative context was the tenth repetition. All of the experimental groups were given the first repetition. E2 was given the first three repetitions, E3 was given the first, ninth, and tenth repetitions, E4 was given the first seven repetitions, and E5 was given all ten repetitions. The context ratings for each repetition are shown in Table 3.

Table 3 Context ratings for each repetition

| Target word | Repetition | | | | | | | | | |
|-------------------|------------|---|---|---|---|---|---|---|---|----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| ancon (hospital) | 4 | 4 | 4 | 4 | 3 | 2 | 2 | 2 | 2 | 2 |
| faddam (write) | 4 | 4 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| masco (train) | 4 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 2 |
| pacon (wear) | 4 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| sagod (visit) | 4 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 1 | 1 |
| denent (remember) | 4 | 4 | 4 | 3 | 3 | 3 | 3 | 3 | 3 | 2 |
| dangy (street) | 3 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 |
| cader (lunch) | 3 | 3 | 3 | 3 | 2 | 2 | 2 | 2 | 2 | 2 |
| hodet (face) | 4 | 4 | 3 | 3 | 3 | 3 | 3 | 2 | 2 | 2 |
| tasper (evening) | 4 | 4 | 3 | 3 | 3 | 3 | 3 | 2 | 2 | 2 |

2.5.15 Word Pairs

In the word pairs treatment, the 10 nonsense words were presented together with the words which they replaced in the sentence contexts. The nonsense words were on the left. The following examples are from the treatment.

| | |
|--------|----------|
| tasper | evening |
| denent | remember |
| ancon | hospital |

All of the word pairs were presented together on one page (see Appendix B). The order of the nonsense words was the same as the order of the nonsense words in E1. The subjects in the word pairs group were not given any instruction on how to learn the word pairs. Therefore, some may have tried to learn them receptively, and others may have tried to learn them productively. The word pairs subjects had the same amount of time (four minutes) as E1 to learn the word pairs.

2.5.16 Time

The time taken to view target words in contexts was controlled. Each group had an equal amount of time per page. Pilot studies were used to determine the length of time necessary to comprehend one page of ten contexts. Four minutes, approximately 24 seconds per context, was found to be sufficient for students to read all of the contexts on a page. It allowed subjects enough time to consider each context but not so much time that they might begin to use mnemonics to remember target words. The time of the treatment varied between groups, as each experimental group read a different number of pages. Experimental group 1 which only read target words once in context completed the treatment in four minutes, whereas, experimental group 5 which read ten contexts for each target word completed the treatment in 40 minutes. The comparison group was given the same amount of time as E1, four minutes, to learn from word pairs.

2.5.17 Dependent Measures

Ten tests were used to measure the effects of incidental learning and learning from word pairs on vocabulary knowledge (see Appendix C). All of the measures were designed to tap into the different aspects of knowledge that may be acquired after meeting a word in context. Each test enabled learners to demonstrate a specific aspect of word knowledge. The tests measured the following:

- Receptive and productive knowledge of orthographic form;
- Receptive knowledge of meaning and form (at two different sensitivities, receptive recognition and receptive recall);
- Receptive and productive knowledge of grammatical functions;
- Receptive and productive knowledge of association;
- Receptive and productive knowledge of syntax.

It was expected that the subjects would score higher on the tests that measured receptive recognition. All of these tests were multiple choice tests. The questions provided subjects with information that could help them select the correct answer. The distractors were not too closely related, and allowed subjects to score correctly with partial knowledge of target words. The tests which measured productive knowledge, or in the case of meaning and form, receptive recall, were much more demanding. None of these tests provided contextual clues to cue recall. Subjects were forced to demonstrate a much fuller knowledge of the target words.

As mentioned above, it was important to have tests at different sensitivities to measure depth of knowledge. Since some subjects in this study met target words very few times, tests that measured partial knowledge were more useful than tests that measured full knowledge. However, as some subjects also met target words ten times dependent measures that assessed a fuller knowledge of the word were also needed.

2.5.18 Productive knowledge of orthographic form

The first type of knowledge that was measured was productive knowledge of orthographic form. This test was completed first because of the likelihood that the subjects would gain orthographic knowledge from the other dependent measures. In this test, the learners heard each target word pronounced twice, they then had 10 seconds to correctly write an item. The amount of time subjects had to write the target words was determined in a pilot study. A native speaker prerecorded the target words. It should be noted that the subjects only met the words in a written context. Use of phonological cues to measure orthographic knowledge not only tested whether the subjects could correctly write the target words but also demonstrated that they could recognize the pronunciation of the word after meeting it solely in a written context. It was found in an earlier pilot study that subjects might have difficulty correctly spelling words containing the letters 'l' and 'r'. This is a problem common to Japanese learners. As a result, the nonsense words were revised with only one of 10 containing those letters.

To score correctly the subjects had to write the target words correctly. If there were any spelling mistakes the answers were marked incorrect. This is because the learners were given the phonological forms of the target words as a cue to recall. Since the subjects were at the intermediate level and were likely to have learned most if not all of the rules of spelling, phonological cues would be enough to at least lead them to write a close approximation of the target words. If responses with minor spelling mistakes were marked as correct, then it could not be determined whether it was due to the learning task or the phonological prompt.

2.5.19 Receptive knowledge of orthographic form

The second dependent measure to be completed was used to assess receptive knowledge of orthographic form. This test was done second because of the likelihood that the subjects would gain orthographic knowledge from the other dependent measures. It should be noted that the first test in which subjects heard each target word pronounced increased the number of exposures to form by one for each group. However, the only test that it could affect was this test.

A multiple-choice format was used with the correct spelling of a target word appearing along with three distractors. The distractors were created to resemble the target words both phonetically and orthographically. Subjects were asked to circle the correctly spelled target words. The following example is for the target word 'denent':

- a) denant b) danant c) danent d) denent

2.5.20 Receptive recall of meaning and form

Receptive knowledge of meaning and form was measured using a receptive recall translation test. On this test a list of the target words was presented with a blank beside each item. The learners had to write the L1 translation to score a correct response.

Having subjects write L1 translations as the response eliminates any misunderstandings which may occur in multiple-choice tests which involve definitions written in the L2. It also reduced the time in which it took students to complete the test. In this experiment there were ten tests so fatigue and time were important factors to consider. Nation (2001) suggests that having learners write L1 translations may be a more effective method of measuring meaning and form than multiple-choice tests. A receptive recall test was used instead of a productive translation test because it was felt that the learners were more likely to write L2 synonyms than the target words on a productive measure. Contextual cues were not provided because subjects could gain knowledge of syntax, and grammatical functions which would aid them in the other tests.

2.5.21 Productive knowledge of grammatical functions

On this test the subjects were asked to write the target words in a sentence. Responses were scored as correct if the target word was written in context with grammatical accuracy. For example, the target word 'ancon' (hospital) would have received scores of '0' for "The girl anconed to school," and '1' for either "It is a beautiful ancon," or "Many doctors and nurses work at ancons." Although the third example shows a clear understanding of the meaning of 'ancon,' points were only awarded for grammatical accuracy. The researcher and another rater scored each response. The inter-rater agreement was 100 percent.

2.5.22 Productive knowledge of syntax

In this test, subjects were asked to write words which had a syntactic relationship with each target word. Words that may often appear in context with a target word were scored as correct. For example, 'letter' and 'diary' were scored as correct for the target word 'faddam' (write). Since subjects often did not understand the instructions during pilot testing, instructions were provided in English and Japanese. Knowledge of syntax could have been measured in the sentences written on the productive knowledge of

grammatical functions text. However, that would have been a very demanding test. Measuring more than one knowledge type in one test may not provide accurate results. For example, a subject who is able to write a target word in a context with grammatical accuracy may not be able to demonstrate another aspect of knowledge such as meaning or syntax. In this situation the subject may write nothing despite having productive knowledge of grammatical functions. Isolating the knowledge type and measuring it in another test may take longer but it is a more effective method of determining gains in knowledge.

2.5.23 Productive knowledge of association

Since productive knowledge of association has only been measured in one of the previous studies investigating learning from context or word pairs (Horst, Cobb, and Meara, 1998) and in that study the definition of association (paradigmatic and syntagmatic associations) differed from the one used in this study (paradigmatic associations), a test had to be designed. The aim was to isolate knowledge of association and determine whether subjects had gained fuller knowledge than on the receptive measure. On this test subjects were presented with the target words and asked to write an associate beside each item. Coordinates, superordinates, subordinates, and synonyms were scored as correct. Antonyms were included in the coordinates category using Aitchison's (1994) definition of coordination. Since this test was designed to measure knowledge of paradigmatic associations, responses that had a syntactic relationship were scored as incorrect. This was explained on the test in both English and Japanese. The following are some of the answers that would be acceptable for the word 'sofa' (the type of association is in brackets): chair (coordinate), chesterfield (synonym), furniture (superordinate), and cushion (subordinate/coordinate).

2.5.24 Receptive knowledge of grammatical functions

Receptive knowledge of grammatical functions was measured using a multiple-choice test. Subjects were presented with three sentences containing the target word one of which was correct. Since the final test measured receptive knowledge of meaning and form, the sentences used in this test could not reveal aspects of the target word's meanings. Knowledge of the target word's part of speech was sufficient for the subjects to choose the correct answer. For example,

- | | |
|-------|--|
| ancon | <ol style="list-style-type: none">1. It is an ancon.2. He anconed.3. It is very ancon. |
|-------|--|

In the above example, knowing that 'ancon' (hospital) is a noun provided subjects with enough information to select the correct answer. Asking learners to simply identify whether the target word was a noun, verb, or adjective may also have been an effective test. However, providing the subjects with the items in context may have aided them in demonstrating partial knowledge. Since the test measuring productive knowledge of grammatical functions was perhaps the most difficult test, it was important to have a receptive test that was sensitive to very small gains.

2.5.25 Receptive knowledge of syntax

A multiple-choice test was used to measure receptive knowledge of syntax. To demonstrate knowledge of syntax, subjects had to select a word which often appears in context with the target word from four choices. The correct answers and all of the distractors occurred in the experimental treatments. The following example is from the receptive knowledge of syntax test (hodet means 'face'):

| | | |
|-------|----|--------|
| hodet | 1. | town |
| | 2. | mirror |
| | 3. | leave |
| | 4. | serve |

Cloze tests would also have been effective in measuring knowledge of syntax. However, the contextual cues may have led to gains in knowledge that would have helped them in the final two tests.

2.5.26 Receptive knowledge of association

A multiple-choice test was used to measure receptive knowledge of association. Subjects had to select the correct response from four choices one of which was linked semantically to the target word. To avoid confusion none of the distractors were frequent collocates of the target words, and the subjects were given the instructions in English and Japanese. The following example was used on the test (denent means 'remember'):

| | | |
|--------|----|---------|
| denent | 1. | buy |
| | 2. | welcome |
| | 3. | forget |
| | 4. | clean |

The correct answer 'forget' is an antonym of the target word and none of the distractors are likely to appear in context regularly with 'remember'.

2.5.27 Receptive recognition of meaning and form

Receptive knowledge of meaning and form was also measured using a multiple-choice recognition test. On this test the target words were presented with the L2 words which had been replaced in the contexts and three distractors. The distractors were of the

same parts of speech as the target words and had meanings that related to contexts read in the treatment. For example, the target word 'ancon' was presented with its real meaning 'hospital' and three other nouns 'house', 'car', and 'city' all of which were mentioned in contexts read in the treatment. This test may only indicate partial knowledge of meaning and form. The distractors could have been much more closely related in meaning. This test was completed last because it was likely to contribute to a learning effect. Other test types such as those that provide contextual cues may also have been effective.

2.6 Case studies

Case studies were done with six learners. Each subject completed one of the treatments from the main study. In the case studies, the subjects were asked to think aloud during both the treatment and tests. After the completion of the tests, questions were asked to gain further insight into the learning process. The following observations emerged from the five incidental reading case studies.

1. All of the subjects attempted to learn some of the contexts by rote. Learning contexts by rote appeared to aid the subjects in gaining knowledge of syntax and grammar. During testing students were often able to recall a context to help them determine an answer. This result may be more applicable to Japanese students because memorization is a common learning strategy (Bamford, 1993).
2. Being unable to produce a correct response on the recall of meaning and form measure did not mean that subjects could not produce a correct response on the productive knowledge of syntax, grammatical functions, and association measures. This appeared to be the result of subjects learning the context by rote and then recalling it during a test.

3. On two occasions subjects were unable to demonstrate knowledge that they had gained until after a test was completed. This suggests that the gains found in the study could be lower than the actual learning gains.
4. All five of the learning from context subjects reported that they would have done much better on the tests if the tenth test, receptive knowledge of meaning and form, had been first. This indicates that most learners had only partial knowledge of meaning and form. They were able to recognize the meanings of the target words but unable to recall the meaning. The multiple-choice recognition test was designed to tap partial knowledge, whereas test three, the recall of meaning and form test, was designed to measure a much fuller knowledge which most subjects could not demonstrate.
5. Subjects often use multiple repetitions to confirm or refute learning. This supports van Daalen-Kaptein's and Elshout-Mohr's (1981) findings that suggest that learners reassess their knowledge about unknown or partially known words from repetition to repetition.
6. There were seven possible outcomes from meeting a word in context. They are as follows:
 1. Learning or partially learning the meaning and form.
 2. Learning or partially learning the meaning and form but later forgetting its form.
 3. Learning or partially learning the meaning and form but later forgetting its meaning.
 4. Building upon previous gains in knowledge of meaning and form.
 5. Reassessing previous gains in knowledge of meaning and form.
 6. Learning an incorrect meaning.
 7. Not understanding the meaning of the word and essentially ignoring its semantic aspect.

7. Forgetting was a very common occurrence which may in part be caused by the experimental design. Learning ten words in a short time may not represent typical incidental learning conditions. However, Nation (2001) suggests that forgetting is a common occurrence when learning from context, and Pressley, Levin, and McDaniel (1987) point out that inferring the meaning of a word from context does not guarantee that it will be remembered.

2.7 Results

The descriptive statistics (means, standard deviations, and number of subjects) of vocabulary knowledge scores for the 10 dependent measures are reported in Table 4. To determine whether there were any overall differences among the treatment groups, a multivariate analysis of variance (MANOVA) was performed using the scores on the 10 dependent measures (productive orthography, receptive orthography, receptive recall of meaning and form, receptive recognition of meaning and form, productive association, receptive association, productive syntax, receptive syntax, productive grammar, and receptive grammar). The independent variable was the type of learning task (incidental vocabulary learning from one, three informative, three less informative, seven, and 10 repetitions, and direct learning from word pairs). An SPSS for MS Windows Release 10.1 package was used for this analysis, and all of the analyses in the experiments in this thesis. The MANOVA realized an overall statistically significant difference (Pillai's Trace ($F(50,675)=3.97$), Hotelling's Trace ($F(50,647)=5.44$), Wilks' Lambda ($F(50,601)=4.67$), and Roy's Greatest Root ($F(10,135)=18.90$) below $p < .000$. A post hoc LSD test was performed in order to locate the sites of significant effects for the aspects of knowledge (see Tables 5-14).

Table 4 Means and standard deviations of learning conditions on dependent measures in Experiment 1

| Condition Contexts | <i>n</i> | PO | RO | RRM | RM | PA | RA | PS | RS | PG | RG |
|-----------------------|----------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 1 repetition | 23 | 4.96 | 6.70 | 0.35 | 5.78 | 0.65 | 4.78 | 0.91 | 4.00 | 1.09 | 5.65 |
| 1 | | (1.94) | (1.72) | (0.65) | (2.09) | (1.03) | (1.76) | (1.62) | (1.62) | (1.62) | (1.67) |
| 3 repetitions | 24 | 5.46 | 7.54 | 0.13 | 4.38 | 0.29 | 3.33 | 0.38 | 4.13 | 0.54 | 5.50 |
| 1, 9, 10 | | (2.17) | (1.96) | (0.34) | (1.74) | (0.62) | (2.04) | (0.77) | (1.54) | (0.88) | (2.00) |
| 3 repetitions | 26 | 5.96 | 8.00 | 1.31 | 6.77 | 2.31 | 5.69 | 1.46 | 5.54 | 2.00 | 6.81 |
| 1-3 | | (1.89) | (1.30) | (1.81) | (2.45) | (2.24) | (2.05) | (1.92) | (2.02) | (2.14) | (1.77) |
| 7 repetitions | 26 | 7.19 | 8.27 | 2.65 | 6.50 | 2.38 | 5.58 | 2.31 | 5.46 | 2.19 | 7.04 |
| 1-7 | | (1.41) | (1.51) | (2.77) | (2.10) | (2.59) | (2.23) | (2.72) | (1.92) | (3.09) | (2.07) |
| 10 repetitions | 24 | 7.71 | 8.75 | 2.88 | 7.58 | 3.96 | 6.88 | 3.67 | 6.63 | 4.25 | 7.96 |
| 1-10 | | (1.94) | (1.03) | (2.40) | (2.12) | (3.33) | (2.40) | (3.46) | (1.84) | (3.91) | (1.97) |
| Word pairs | 23 | 7.04 | 8.09 | 6.22 | 9.57 | 4.78 | 7.04 | 4.74 | 6.35 | 5.57 | 7.09 |
| | | (2.18) | (1.56) | (2.66) | (0.79) | (3.04) | (2.12) | (2.68) | (2.50) | (3.20) | (2.00) |
| Control | 22 | 1.35 | | | | | | | | | |
| no repetitions | | (1.02) | | | | | | | | | |

Note. Maximum score = 10, standard deviations are in parentheses, Key. PO=productive knowledge of orthography, RO=receptive knowledge of orthography, PM=productive knowledge of meaning and form, RM=receptive knowledge of meaning and form, PA=productive knowledge of association, RA=receptive knowledge of association, PS=productive knowledge of syntax, RS=receptive knowledge of syntax, PG=productive knowledge of grammar, RG=receptive knowledge of grammar.

Table 4 shows that gains in all aspects of knowledge tended to increase as the number of repetitions increased. The results show that when time spent on the treatments was equal, subjects that completed the learning from word pairs task performed better on all 10 tests than those that did the single repetition incidental learning condition.

Moreover, the learning from word pairs task contributed to larger gains than the ten repetition group on six of the 10 measures despite taking 10 percent of the time. The quality of the context was also found to affect learning. Mean scores were higher on all ten measures when subjects met the target words in three more informative contexts than in three less informative contexts.

2.7.1 Productive knowledge of orthography

The LSD test showed that learning productive knowledge of orthography from word pairs was a more effective task than incidental learning. The single repetition group and the word pairs group were given the same amount of time (four minutes) to learn the target words; however, the word pair group gained significantly more knowledge

($p<.000$). The word pairs group also gained more knowledge than the three less informative repetitions group ($p<.01$). The results also showed that learners gained significantly more productive knowledge of orthography as repetitions increased. Learners in the seven ($p<.000$) and 10 ($p<.000$) repetition groups gained significantly more knowledge of orthography than the single repetition group. The three informative repetitions group was also outperformed by the seven ($p<.05$) and 10 ($p<.01$) repetition groups. Since there were no significant differences found between the three informative and three less informative repetitions groups, the quality of context was not found to affect productive knowledge of orthography.

Table 5 LSD test of differences in productive knowledge of orthography among the six groups

| | Group | | | | | |
|----------------|--------|---------------------|--------------------|---------|----------|------------|
| | 1 rep. | 3 informative reps. | 3 less informative | 7 reps. | 10 reps. | Word pairs |
| 1 repetition | - | NS | NS | *** | *** | *** |
| 3 informative | | - | NS | * | ** | NS |
| 3 less info. | | | - | ** | *** | ** |
| 7 repetitions | | | | - | NS | NS |
| 10 repetitions | | | | | - | NS |
| Word pairs | | | | | | - |

Note. * $p<.05$, ** $p<.01$, *** $p<.000$

A one way analysis of variance (ANOVA) was performed using the scores for the single repetition group and the control group from the test of productive knowledge of orthography. The independent variable was the number of repetitions, zero for the control group and 1 for the single repetition group. The ANOVA realized a statistically significant difference ($F(1,44) = 62.06, p<.000$). This result indicates that the gains in productive knowledge of orthography for all of the experimental groups and the comparison group can be attributed to the treatments rather than the phonological cue on the test.

2.7.2 Receptive knowledge of orthography

The LSD test showed that learning receptive knowledge of orthography from word pairs was a more effective task than incidental learning. The single repetition group and the word pair group were given the same amount of time (four minutes) to learn the target words; however, the word pair group gained significantly more knowledge ($p < .01$). The results also showed that learners gained significantly more receptive knowledge of orthography as repetitions increased. After three informative ($p < .01$), seven ($p < .000$), and 10 ($p < .000$) repetitions learners demonstrated larger gains in knowledge than after one repetition. The 10 repetition group also outperformed the three less informative repetitions group ($p < .01$). Quality of context was not found to affect receptive knowledge of orthography. Although scores were higher for the three informative repetitions group than the three less informative repetitions group, there was not a significant difference between the scores.

Table 6 LSD test of differences in receptive knowledge of orthography among the six groups

| | Group | | | | | |
|----------------|--------|---------------------|--------------------|---------|----------|------------|
| | 1 rep. | 3 informative reps. | 3 less informative | 7 reps. | 10 reps. | Word pairs |
| 1 repetition | - | ** | NS | *** | *** | ** |
| 3 informative | | - | NS | NS | NS | NS |
| 3 less info. | | | - | NS | ** | NS |
| 7 repetitions | | | | - | NS | NS |
| 10 repetitions | | | | | - | NS |
| Word pairs | | | | | | - |

Note. * $p < .05$, ** $p < .01$, *** $p < .000$

2.7.3 Receptive recall of meaning and form

The LSD test showed that receptive recall of meaning and form from word pairs was a much more effective task than incidental learning. The word pairs group demonstrated larger gains than all of the incidental groups on this test ($p < .000$). The results also

showed that learners gained significantly more productive knowledge of meaning and form as repetitions increased. The 10 repetition group outperformed the three informative ($p<.01$), three less informative ($p<.000$), and single ($p<.000$) repetition groups. The seven repetition group also outperformed the three informative ($p<.05$), three less informative ($p<.000$), and single ($p<.000$) repetition groups. The results also showed that the quality of context had a significant effect on gaining receptive recall of meaning and form. The three informative repetitions group demonstrated larger gains than the three less informative repetitions group ($p<.05$).

Table 7 LSD test of differences in receptive recall of meaning and form among the six groups

| | Group | | | | | |
|----------------|------------|------------------------|-----------------------|------------|-------------|------------|
| | 1 reps. | 3 informative reps. | 3 less informative | 7 reps. | 10 reps. | Word pairs |
| 1 repetition | - | NS | NS | *** | *** | *** |
| 3 informative | | - | * | * | ** | *** |
| 3 less info. | | | - | *** | *** | *** |
| 7 repetitions | | | | - | NS | *** |
| 10 repetitions | | | | | - | *** |
| Word pairs | | | | | | - |

Note. * $p<.05$, ** $p<.01$, *** $p<.000$

2.7.4 Receptive knowledge of meaning and form

The LSD test showed that learning receptive knowledge of meaning and form from word pairs was a much more effective task than incidental learning. The word pairs group outperformed the single ($p<.000$), three informative ($p<.000$), three less informative ($p<.000$), seven ($p<.000$), and 10 ($p<.01$) repetition groups. The results also showed that learners gained significantly more receptive knowledge of meaning and form as repetitions increased. After 10 repetitions learners gained more knowledge than after one ($p<.01$), and three less informative ($p<.000$) repetitions. Also the seven ($p<.000$) repetition group outperformed the three less informative repetitions group. The results also showed that the quality of context had a significant effect on gaining

productive knowledge of meaning and form. The three informative ($p<.000$), and single ($p<.05$) repetition groups outperformed the three less informative repetitions group.

Table 8 LSD test of differences in receptive knowledge of meaning and form among the six groups

| | Group | | | | | |
|----------------|------------|------------------------|-----------------------|------------|-------------|------------|
| | 1 reps. | 3 informative reps. | 3 less informative | 7 reps. | 10 reps. | Word pairs |
| 1 repetition | - | NS | * | NS | ** | *** |
| 3 informative | | - | *** | NS | NS | *** |
| 3 less info. | | | - | *** | *** | *** |
| 7 repetitions | | | | - | NS | *** |
| 10 repetitions | | | | | - | ** |
| Word pairs | | | | | | - |

Note. * $p<.05$, ** $p<.01$, *** $p<.000$

2.7.5 Productive knowledge of association

The LSD test showed that learning productive knowledge of association from word pairs was a much more effective task than incidental learning from reading. The word pairs group demonstrated larger gains than the single ($p<.000$), three informative ($p<.05$), three less informative ($p<.000$), and seven repetition ($p<.05$) groups. The results also showed that learners gained significantly more productive knowledge of association as repetitions increased. The 10 repetition group had larger gains than the single ($p<.01$), three informative ($p<.05$), three less informative ($p<.000$), and seven repetition ($p<.05$) groups. The seven repetition group also outperformed the three less informative repetitions group. The results showed that the quality of context had a significant effect on gaining productive knowledge of association. Learners from the three less informative context groups had smaller gains than the single ($p<.05$), three informative ($p<.000$), and seven ($p<.000$) repetition groups.

Table 9 LSD test of differences in productive knowledge of association among the six groups

| | Group | | | | | |
|----------------|------------|------------------------|-----------------------|------------|-------------|------------|
| | 1 reps. | 3 informative reps. | 3 less informative | 7 reps. | 10 reps. | Word pairs |
| 1 repetition | - | * | NS | * | *** | *** |
| 3 informative | | - | ** | NS | * | *** |
| 3 less info. | | | - | ** | *** | *** |
| 7 repetitions | | | | - | * | ** |
| 10 repetitions | | | | | - | NS |
| Word pairs | | | | | | - |

Note. * $p < .05$, ** $p < .01$, *** $p < .000$

2.7.6 Receptive knowledge of association

The LSD test showed that learning receptive knowledge of association from word pairs was a much more effective task than incidental learning. The word pairs group demonstrated larger gains than the single ($p < .000$), three informative ($p < .05$), three less informative ($p < .000$), and seven repetitions groups ($p < .05$). The results also showed that learners gained significantly more receptive knowledge of association as repetitions increased. The 10 repetition group had larger gains than the single ($p < .01$), three informative ($p < .05$), three less informative ($p < .000$), and seven repetitions groups ($p < .05$). The seven repetition group outperformed the three less informative repetitions group ($p < .000$). The results showed that the quality of context had a significant effect on gaining receptive knowledge of association. The three informative ($p < .000$), and single ($p < .05$) repetition groups outperformed the three less informative repetition group.

Table 10 LSD test of differences in receptive knowledge of association among the six groups

| | Group | | | | | |
|----------------|------------|------------------------|-----------------------|------------|-------------|------------|
| | 1 reps. | 3 informative reps. | 3 less informative | 7 reps. | 10 reps. | Word pairs |
| 1 repetition | - | NS | * | NS | ** | *** |
| 3 informative | | - | *** | NS | * | * |
| 3 less info. | | | - | *** | *** | *** |
| 7 repetitions | | | | - | * | * |
| 10 repetitions | | | | | - | NS |
| Word pairs | | | | | | - |

Note. * $p < .05$, ** $p < .01$, *** $p < .000$

2.7.7 Productive knowledge of syntax

The LSD test showed that learning productive knowledge of syntax from word pairs was a much more effective task than incidental learning. The word pair group demonstrated larger gains than the single, three informative, three less informative, and seven repetition groups ($p < .000$). The results also showed that learners gained significantly more productive knowledge of syntax as repetitions increased. After 10 repetitions learners gained more knowledge than after seven ($p < .05$), three informative ($p < .01$), three less informative ($p < .000$), and one ($p < .000$) repetition. The seven repetition group outperformed the three less informative ($p < .01$), and the single ($p < .05$) repetition groups. Although scores were higher for the single and three informative repetitions groups than the three less informative repetitions group, there was no significant difference between the scores suggesting that the quality of context may not affect productive knowledge of syntax.

Table 11 LSD test of differences in productive knowledge of syntax among the six groups

| | Group | | | | | |
|----------------|------------|------------------------|-----------------------|------------|-------------|------------|
| | 1 reps. | 3 informative reps. | 3 less informative | 7 reps. | 10 reps. | Word pairs |
| 1 repetition | - | NS | NS | * | *** | *** |
| 3 informative | | - | NS | NS | ** | *** |
| 3 less info. | | | - | ** | *** | *** |
| 7 repetitions | | | | - | * | *** |
| 10 repetitions | | | | | - | NS |
| Word pairs | | | | | | - |

Note. * $p < .05$, ** $p < .01$, *** $p < .000$

2.7.8 Receptive knowledge of syntax

The LSD test showed that learning receptive knowledge of syntax from word pairs was a much more effective task than incidental learning. The word pair group demonstrated larger gains than the three less informative ($p < .000$), and single ($p < .000$) repetition groups. The results also showed that learners gained significantly more receptive knowledge of syntax as repetitions increased. After 10 repetitions learners had gained more knowledge than after seven ($p < .05$), three informative ($p < .05$), three less informative ($p < .000$), and one ($p < .000$) repetition. The seven repetition learners also demonstrated more knowledge than the three less informative ($p < .05$), and single ($p < .01$) repetition learners. The three informative repetitions learners also had higher scores than the single repetition group ($p < .01$). Since the three informative repetitions group had higher mean scores than the three less informative repetitions group, the quality of context was found to have a significant effect on gaining receptive knowledge of syntax.

Table 12 LSD test of differences in receptive knowledge of syntax among the six groups

| | Group | | | | | |
|----------------|------------|------------------------|-----------------------|------------|-------------|------------|
| | 1 reps. | 3 informative reps. | 3 less informative | 7 reps. | 10 reps. | Word pairs |
| 1 repetition | - | ** | NS | ** | *** | *** |
| 3 informative | | - | * | NS | * | NS |
| 3 less info. | | | - | * | *** | *** |
| 7 repetitions | | | | - | * | NS |
| 10 repetitions | | | | | - | NS |
| Word pairs | | | | | | - |

Note. * $p < .05$, ** $p < .01$, *** $p < .000$

2.7.9 Productive knowledge of grammatical functions

The LSD test showed that learning productive knowledge of grammatical functions from word pairs was a much more effective task than incidental learning. The word pair group demonstrated greater gains in knowledge than the single ($p < .000$), three informative ($p < .000$), three less informative ($p < .000$), and seven ($p < .000$) repetitions groups. The results also showed that learners gained significantly more productive knowledge as repetitions increased. The 10 repetition group gained more knowledge than the single ($p < .000$), three informative ($p < .01$), three less informative ($p < .000$), and seven ($p < .01$) repetition groups. Also the seven repetitions learners outperformed the three less informative repetitions group ($p < .05$). Although scores were higher for the single and three informative repetitions groups than the three less informative repetitions group, there was no significant difference between the scores suggesting that the quality of context may not affect gaining productive knowledge of grammatical functions.

Table 13 LSD test of differences in productive knowledge of grammatical functions among the six groups

| | Group | | | | | Word pairs |
|----------------|------------|------------------------|-----------------------|------------|-------------|------------|
| | 1 reps. | 3 informative reps. | 3 less informative | 7 reps. | 10 reps. | |
| 1 repetition | - | NS | NS | NS | *** | *** |
| 3 informative | | - | NS | NS | ** | *** |
| 3 less info. | | | - | * | *** | *** |
| 7 repetitions | | | | - | ** | *** |
| 10 repetitions | | | | | - | NS |
| Word pairs | | | | | | - |

Note. * $p < .05$, ** $p < .01$, *** $p < .000$

2.7.10 Receptive knowledge of grammatical functions

The LSD test showed that learning receptive knowledge of grammatical functions from word pairs was a more effective task than incidental learning. The word pairs group gained more knowledge than the single ($p < .05$), and three less informative repetition ($p < .01$) groups. The results also showed that learners gained significantly more receptive knowledge of grammatical functions as repetitions increased. The ten repetition group outperformed the single ($p < .000$), three informative ($p < .05$), and three less informative repetition ($p < .000$) groups. After seven repetitions, learners gained more knowledge than after one ($p < .05$), and three less informative repetitions ($p < .01$). The three informative repetition group also demonstrated more knowledge than the single repetition group ($p < .05$). Quality of context was found to have a significant effect on gaining receptive knowledge of grammatical functions. Learners who met words in three informative contexts demonstrated more knowledge than those who met words in three less informative contexts ($p < .05$).

Table 14 LSD test of differences in receptive knowledge of grammatical functions among the six groups

| | Group | | | | | |
|----------------|------------|------------------------|-----------------------|------------|------------|---------------------|
| | 1 reps. | 3 informative reps. | 3 less informative | 7 reps. | 10 reps | Word pairs reps. |
| 1 repetition | - | * | NS | * | *** | * |
| 3 informative | | - | * | NS | * | NS |
| 3 less info. | | | - | ** | *** | ** |
| 7 repetitions | | | | - | NS | NS |
| 10 repetitions | | | | | - | NS |
| Word pairs | | | | | | - |

Note. * $p < .05$, ** $p < .01$, *** $p < .000$

2.8 Discussion

The results of the present study confirm that Japanese intermediate EFL learners can incidentally acquire vocabulary from reading, thus supporting the findings of previous investigations (Day, Omura, and Hiramatsu, 1991; Dupuy and Krashen, 1993; Horst, Cobb, and Meara, 1998; Hulstijn, 1992; Hulstijn, Hollander, and Greidanus, 1996; Nagy, Herman, and Anderson, 1985; Pitts, White, and Krashen, 1989). The experimental design expanded upon earlier methodologies measuring vocabulary gains in knowledge of orthography, association, syntax, meaning and form, and grammatical functions. The vast majority of incidental vocabulary learning studies have equated gains of meaning and form with acquisition.

Two variables, context and repetition, were controlled to examine how they influence vocabulary acquisition. Repetition was found to have a significant effect supporting earlier findings (Jenkins, Stein, and Wysocki, 1984; Horst, Cobb, and Meara, 1998; Rott, 1999; Saragi, Nation, and Meister, 1978). Gains in knowledge of all five aspects tended to be larger with increased repetitions of target words depending on the degree of information found in the contexts (The issue of context is addressed below).

At one repetition, sizeable gains in receptive knowledge were found for orthography (67%), syntax (40%), association (48%), meaning and form (58%), and grammatical functions (57%). The subjects also demonstrated large gains in productive knowledge of orthography (50%) indicating that for intermediate learners spelling is likely to be the first knowledge type acquired as Schmitt and McCarthy (1997), and Schmitt (1998, 2000) suggest. Productive gains ranged from 3.5% for meaning and form to 50% for orthographic knowledge. Since partial gains were made for all aspects despite relatively little knowledge of meaning, acquiring knowledge of association, syntax, spelling, and grammatical functions may not be dependent on the acquisition of meaning. This is certainly the case in L1. Adult native speakers occasionally demonstrate productive knowledge of a word with grammatical, orthographic, and syntactic accuracy but without accuracy of meaning.

At three repetitions in informative contexts the subjects demonstrated increased learning on every measure. Receptive knowledge of orthography, grammatical functions and syntax, and productive knowledge of association were found to be significantly larger than at one repetition. This suggests that knowledge of those aspects may develop earlier than meaning. Relatively large gains in productive knowledge for each aspect began to be found at this point indicating learners may have gained enough word knowledge to correctly use a target word in context. The gains were 13% (meaning and form), 15% (syntax), 20% (grammatical functions), 23% (association), and 60% (orthography).

At seven repetitions there was relatively little difference between the mean scores at three repetitions with the receptive measures, but larger gains in productive knowledge were demonstrated for all aspects. Productive knowledge of orthography, meaning and form were significantly larger than at three repetitions, as well as syntax, and association at one repetition. This highlights the importance of using multiple tests at different sensitivities to measure gains in knowledge. The differences between E2 (three repetitions) and E4 (seven repetitions) on the meaning and form tests show this clearly. From the scores on the receptive test it appears that there is little difference

between the groups, E4 had a mean score (6.5) slightly lower than E2 (6.8). This suggests that the E4 subjects did not gain any semantic knowledge from the 4th, 5th, 6th, and 7th repetitions. The results of the productive test demonstrated that this was incorrect. E4's mean score (2.65) was more than double that of E2 (1.31). If only the receptive test had been used to measure meaning and form, a repetition effect would not have been found. This occurred four times in the results, and suggests that studies that measure an aspect of knowledge with only one test are unlikely to reveal the full extent of learning. The gains in productive knowledge at seven repetitions were 22% (grammatical functions), 23% (syntax), 24% (association), 27% (meaning and form), and 72% (orthography).

At ten repetitions the subjects were able to demonstrate sizeable learning gains for both productive and receptive knowledge for every aspect with large increases on most of the measures. Receptive orthographic knowledge was 88%, while the other aspects ranged from 66% (syntax) to 80% (grammatical functions) on the receptive measures. The gains on the productive tests were 29% (meaning and form), 37% (syntax), 40% (association), 43% (grammatical functions), and 77% (orthography). The gains were significantly larger than at one repetition for all aspects of knowledge. They were also significantly greater than those at three repetitions for each aspect except for receptive knowledge of meaning and form, and orthography. Receptive knowledge of syntax, and association, as well as productive knowledge of grammar, syntax, and association were also significantly larger than the gains made by the seven repetition group. Since gains continued to be made at ten repetitions, it appears that in many cases more than ten repetitions are needed to achieve full knowledge of a word. By ten meetings with a word, there is the possibility that learners will be able to recognize its spelling, meaning, part of speech, words that it is associated with, and words that it occurs with in context. There is also the possibility that they will be able to use it with grammatical accuracy in a sentence with words that it has a syntactic relationship, be able to spell it correctly, write an association, and write its meaning.

Table 15 Correlation coefficients between the number of repetitions and dependent measures in Experiment 1

| | Productive | | | | | Receptive | | | | |
|------------|------------|-----|-----|-----|-----|-----------|-----|-----|-----|-----|
| | PO | PM | PA | PS | PG | RO | RM | RA | RS | RG |
| repetition | .51 | .52 | .48 | .45 | .44 | .38 | .32 | .37 | .42 | .40 |

Note. All correlations are significant at 0.01 (one-tailed test), Key. PO=productive knowledge of orthography, RO=receptive knowledge of orthography, PM=productive knowledge of meaning and form, RM=receptive knowledge of meaning and form, PA=productive knowledge of association, RA=receptive knowledge of association, PS=productive knowledge of syntax, RS=receptive knowledge of syntax, PG=productive knowledge of grammar, RG=receptive knowledge of grammar.

The correlations between the number of repetitions and relative learning gains for each aspect of knowledge are shown in Table 15. The size of the correlations demonstrates that repetition does play a significant role in gaining vocabulary knowledge supporting the earlier findings by Horst, Cobb, and Meara (1998) and Saragi, Nation, and Meister (1978). Table 15 also shows that repetition correlates more strongly with the productive measures than the receptive measures. This is not surprising since the receptive measures were designed to be sensitive to small gains of knowledge. They may therefore be more sensitive to the influence of other factors. The correlations may also reflect the fact that the receptive tests provided an opportunity to score correctly through guessing while the productive tests did not. The productive tests may therefore provide a more accurate assessment of correlation. The difference between the correlations for the more and the less sensitive tests suggests that there are numerous factors that may influence vocabulary learning.

The second variable that was controlled in this study and was found to have a significant effect on acquisition was context. This supports Beck, McKeown, and McCaslin's (1983) claims, as well as previous findings (Jenkins, Matlock, and Slocumb, 1989). Learners who met target words in the three most informative contexts (E2) produced significantly larger gains in receptive knowledge of meaning and form, association, syntax, grammatical functions, and productive knowledge of association than learners who met the target words in the most informative context, followed by the

two least informative contexts (E3). Knowledge of orthography was also greater for E2; however, the differences were not significant. Furthermore, learners in the one repetition group (E1) significantly outperformed E3 on two measures, receptive knowledge of association, and meaning and form, and had higher mean scores on seven of the ten measures. Orthography, and receptive knowledge of syntax, and grammatical functions were also affected by context but not as profoundly as the other knowledge types. The relatively tiny gains made by E3 in relation to E1 and E2 suggest that initial gains resulting from meeting an unknown word in one informative context may be reduced in subsequent meetings if those contexts are less informative or misleading. It is likely that meeting a partially known word in an uninformative context may lead learners to reassess their previous knowledge or forget what they had learned. However, this may depend on the amount of knowledge previously gained. Although the final three contexts met by the ten repetition group (E5) were rated as the least informative (see Table 3), E5 produced larger gains than the seven repetition group (E4) on all ten measures. This suggests that there may be a knowledge threshold after which backsliding is less likely to occur. If there is indeed a knowledge threshold, determining the point at which backsliding ceases to occur would be a very useful finding. It could provide researchers with a quantifiable amount that indicates whether acquisition is likely to occur. Further research is also necessary to determine which aspects of knowledge are gained from meeting an unknown word in different sequences of informative and uninformative contexts. The order of contexts in this study always began with the most informative contexts. However, learners may initially meet unknown words in less informative contexts when reading. Meeting an unknown word in a less informative context is a very common occurrence that needs to be examined.

Nation (1982) pointed out that unless there is a clear definition of context in research, results may be misleading. The sizeable context effect found in this study reinforces this and indicates that a better understanding of context is needed. Since there are so many different types of contexts used in incidental learning research, it is very unlikely that the results translate from one context to another. Certainly gains in knowledge are

less likely to occur from reading a graded reader than from a passage created with sufficient clues to infer the meaning of unknown words.

The results of the present study support Nation and Wang's (1999: 363) claim that "there is no set number of repetitions that will ensure learning." It would appear that the effects of context and other variables have a very strong influence on the number of repetitions needed to gain knowledge of a word. Further research investigating context and repetition effects is necessary to determine how they affect vocabulary learning. The findings do indicate that if learners meet unknown words a minimum of ten times during reading, sizeable vocabulary growth may occur. Studies by Wodinsky and Nation (1988), and Nation and Wang (1999) indicate that in an extensive reading program with graded readers there may be sufficient repetitions to promote significant vocabulary growth.

The results of the present study also indicate that learning from word pairs is a very effective method of gaining vocabulary knowledge supporting earlier findings (Laufer and Shmueli, 1997; Thorndike, 1908; Webb, 1962). The group which was given the ten target words together with their L2 synonyms (WP) demonstrated large gains for all five knowledge types, and outperformed E5, the ten repetition group, on six of the ten measures. However, only the scores on the measures of meaning and form were significantly greater than those of E5. E5 produced larger gains in receptive knowledge of grammatical functions, syntax, and productive and receptive knowledge of orthography. This suggests that in the long run incidental learning from reading may be a more effective method of gaining grammatical, syntactic, and orthographic knowledge than word pairs. However, the large gains in every knowledge type made by WP indicate that learning from word pairs is an effective way of gaining not only knowledge of meaning and form, but also orthography, syntax, grammatical functions, and association. The word pair group significantly outperformed E1 on all of the tests, and E2 and E4 on six of the tests. This suggests that if there is a limited time available for vocabulary learning, learning from word pairs is superior to incidental learning. Surprisingly, gains in syntax, and grammatical functions were similar to the 10

repetition group. This is despite the fact that meeting an unknown word in context would seem to be a much better way of learning its grammatical and syntactic properties than out of context and paired with a synonym. The reason for such large gains in those knowledge types from word pairs may be that the subjects made assumptions about the target words based on their L2 knowledge of their synonyms.

All of the target word synonyms used in the word pairs were high frequency words for which the subjects were very likely to have had full knowledge. Since they were not told that the target words were nonsense words before the experiment, it is most likely that they believed they were unknown less frequent synonyms. For example, for the target word *ancon* they were given the meaning *hospital*. In this case they might assume *ancon* is a less frequent synonym such as *clinic* or *surgery*. Since they were likely to have syntactic and grammatical knowledge of *hospital*, they could use it to answer questions about *ancon*. Transfer of knowledge from more frequent synonyms to less frequent synonyms may help to explain how less frequent vocabulary is acquired relatively quickly. If a learner realizes after one or two occurrences that an unknown word has a more frequent synonym, a transfer of knowledge may speed up the learning process. The learning burden (Nation, 1990, 2001) would be less, and therefore, acquisition should be faster. This suggests that different words are acquired at different rates. Further research examining what is gained from word pairs with different learning burdens may show that some L2 words are more difficult to learn than others. For example, a nonsense word paired with a low frequency word that does not have a high frequency synonym may produce much lower gains in knowledge of syntax, and grammatical functions because learners have never seen a similar word in context. Determining whether L2 words have different learning burdens would be useful from a pedagogical standpoint because it would help teachers ascertain which words may need explicit instruction and which words do not.

2.9 Summary of main findings

The findings of this study suggest:

1. Japanese intermediate EFL learners may acquire knowledge of orthography, meaning and form, association, syntax, and grammatical functions through incidental vocabulary learning from reading.
2. Repetition affects incidental vocabulary learning from reading. Learners who encounter an unknown word more times in informative contexts are able to demonstrate significantly larger gains in vocabulary knowledge types than learners who have fewer encounters with an unknown word.
3. If learners meet unknown words ten times in context sizeable learning gains may occur. However, to develop full knowledge of a word more than ten repetitions may be needed.
4. The amount of information in a context strongly influences the size of gains in meaning and form, association, syntax, and grammatical functions. Partial gains from an informative context may be reduced or eliminated if followed by uninformative or misleading contexts.
5. There may be a knowledge threshold after which backsliding is less likely to occur. Finding it would provide a level of knowledge that indicates whether acquisition will be successful.
6. Learning vocabulary from word pairs is an effective time-efficient method of acquiring vocabulary. It promotes not only large gains in knowledge of meaning and form but also association, syntax, grammatical functions, and orthography.
7. Both learning from word pairs and incidental vocabulary learning from reading are effective methods for gaining significant knowledge of a word. However, unless learners read extensively the results suggest that word pairs may be a more useful approach.

8. Using only one test to measure an aspect of knowledge is unlikely to reveal the full extent of vocabulary learning that has occurred. Multiple tests of a knowledge type are necessary to capture full vocabulary gains.

2.10 Limitations

It is important to note that the learning conditions in the present study do not mirror those of incidental learning. Incidental learning from reading is most likely to occur when reading a number of texts containing repeated occurrences of unknown words. Controlled variables in this study such as context, repetition, and the number of unknown words will vary from text to text, and day to day. Further research is needed to confirm and expand upon the findings. Investigating how repetition and context influence incidental learning under more ecologically valid reading conditions would be a useful follow up to this study.

Furthermore, the interpretation of the results is limited by the age and proficiency of the subjects. Learner factors have been found to influence incidental learning (e.g. Jenkins, Stein, and Wysocki, 1984; McKeown, 1985; Sternberg and Powell, 1983). A replication of this study with EFL learners of differing ages and proficiencies may improve our understanding of how context, repetition, and meaning affect the different aspects of word knowledge.

A third limitation of the present study is that the subjects are likely to have had full knowledge of L2 synonyms for each target word. Subjects in the study were unaware that the target words were not real English words. It is likely that when they gained knowledge of the meaning of a target word, they believed it was a less frequent L2 synonym of a word that they knew. It was postulated that gaining syntactic, associative, and grammatical knowledge of an L2 synonym might be easier than learning an entirely unknown word because of the reduced learning burden. If this is correct, then the gains in knowledge of syntax, grammatical functions, and association

found in this study may be higher than for words with a greater learning burden. Future research needs to examine whether different words are acquired at different rates. A follow up study that used different types of words may help to clarify this issue.

2.11 Rationale for the next experiment

One of the more surprising findings of the previous study was that students learning from word pairs demonstrated larger gains in productive knowledge of grammatical functions and syntax than students who met each target word in ten different contexts. Learning a word in context would seem to be a much better way of gaining its syntactic and grammatical properties. It was hypothesized that the subjects gained knowledge of the unknown target words from their known synonyms. If this were true then different words would be learned at different rates. A nonsense word paired with a L1 translation which had a known L2 synonym may be acquired more quickly because of the reduced learning burden than a nonsense word paired with a L1 translation for which a learner had no L2 knowledge. When a learner acquires the meaning and form of a less frequent synonym, they may gain more than just knowledge of meaning and form. If they have full knowledge of the more frequent synonym, they may use that as a model for the less frequent word. They may make assumptions about syntax, associations, and grammatical functions based on what they know about the synonym. These aspects of knowledge may be acquired extremely quickly. To determine this, a second word pair experiment needs to be completed. In this experiment, L2 word knowledge will be a variable. Ten L1 words that have known L2 synonyms will be paired with nonsense target words, and ten L1 words that do not have known L2 synonyms will also be paired with nonsense words. The 20 target words will also be learned in contexts with their L1 meanings given to investigate how context affects learning. Tests measuring receptive and productive knowledge of orthography, association, meaning and form, syntax, and grammatical functions will be the dependent measures. This study may be useful for several reasons. First, it may help researchers to better understand the vocabulary acquisition process. Incidental

learning theory suggests that young L1 learners acquire large amounts of vocabulary very quickly through repeated meetings with words in contexts (Nagy, Herman, and Anderson, 1985). Gaining knowledge of less frequent words from more frequent synonyms would help to explain the rapid vocabulary growth that occurs with young L1 learners. Second, determining whether words have varying rates of acquisition would be useful for teachers and learners. They could develop teaching and learning strategies based on a word's learning burden.

CHAPTER 3

3.0 Introduction

Decontextualized vocabulary learning tasks such as learning from word pairs have consistently been found to contribute to large gains in knowledge of meaning and form (see for example, Thorndike, 1908; Webb, 1962). This is not surprising because learning from word pairs provides learners with the meanings of target words, and focuses learners solely on meaning and form. In contrast, contextualized tasks such as incidental vocabulary learning from reading may focus learners on many different aspects of vocabulary knowledge. Unsurprisingly, comparisons between incidental vocabulary learning and learning from word pairs have shown that much larger gains in knowledge of meaning and form are made by learners who took part in the decontextualized tasks (Laufer and Shmueli, 1997; Prince, 1996; Seibert, 1930). Unfortunately, none of those studies measured any other aspects of knowledge such as syntax or grammatical functions that seem more likely to be gained through contextualized learning.

Experiment 1 may be the only study that has compared the two tasks and measured several aspects of knowledge. The results showed that subjects who had completed the decontextualized task gained significantly more knowledge of grammatical functions, syntax, association, orthography, and meaning and form than those who had learned in context for the same amount of time. It is surprising that learning in context did not contribute to larger gains in knowledge of syntax and grammatical functions because those aspects appear to be ignored in decontextualized tasks. One explanation that was given in Experiment 1 was that the subjects might have used their knowledge of the target word's L2 synonyms in the learning from word pairs task to gain syntactic and grammatical knowledge of the target words. If learners gain knowledge of meaning and form for a word that has a more frequent synonym which they know, they may be able to use the known synonym as a model. They may then make assumptions about

syntax and grammar for the less frequent synonym based upon what they know about the more frequent synonym. However, since Experiment 1 was designed to measure the effects of repetition on incidental learning rather than contextualized vs. decontextualized learning, confounding variables made it difficult to determine why the learning from word pairs task was superior. Comparing learning from word pairs and learning in glossed sentences would provide much more accurate data on the effects of context on the different aspects of knowledge.

The present study investigated the effects of context and synonymy on vocabulary learning. It compared learning from word pairs and learning from glossed sentences, and learning synonyms and non-synonyms. Specifically, it examined the following:

1. How does context affect vocabulary learning?
2. Which is more effective, learning from word pairs or learning from glossed sentences?
3. Which aspects of knowledge are gained through learning from word pairs and learning from glossed sentences?
4. How does synonymy affect vocabulary learning?

3.1 Research investigating the effects of context on vocabulary learning

Many studies have investigated learning in context and decontextualized vocabulary learning (Dempster, 1987; Griffin 1992; Laufer and Shmueli 1997; Pickering, 1982; Prince, 1996; Seibert, 1930). In most of those studies, incidental learning from context and explicit learning in a decontextualized task were compared making it impossible to accurately assess the effects of context. However, there are four studies that have examined explicit learning from context and explicit learning from word pairs (Dempster, 1987; Griffin, 1992; Laufer and Shmueli, 1997; Seibert, 1930).

In a series of L1 experiments, Dempster (1987) compared learning from word pairs, learning from a single sentence, and learning from three sentences. In all three conditions, the meanings of the target words were provided. A receptive recall measure was administered after the treatments. In the test, subjects were given the target words and had to write their meanings. Dempster did not find any significant differences between the three conditions. Since the recall measure may have favored the learning from word pairs condition, two productive measures incorporating context were added in the final experiment. No significant differences were found on cued recall and sentence production tests. One reason why significant differences were not found in any of the experiments may be that all of the tests used in the study measured relatively full knowledge of meaning and form. Less demanding measures such as receptive recognition tests may have produced very different results. On the sentence production test, subjects had to demonstrate that they knew the meaning of the target words by using them in sentences. That would appear to be a much more demanding measure than the earlier test of receptive recall. Of the 1368 possible responses, the subjects only completed fifteen responses, and only five of them were scored as correct. Whether those sentences demonstrated knowledge of grammatical functions and syntax was not discussed. Perhaps, if the sentence production test had measured another aspect of knowledge such as grammatical functions, or collocation, subjects in the sentence learning condition may have outperformed learners in the word pair group. In this regard, the results of Experiment 1 supported Nagy, Herman, and Anderson's (1985) claim that multiple tests at different sensitivities are needed to determine how much learning has occurred.

Griffin (1992) compared learning 20 vocabulary items from word pairs and sentences with higher and lower ability foreign language learners. In the sentence learning condition, translations of the sentences were written above each context. Productive and receptive knowledge of meaning and form were measured with cued recall tests. In the tests, subjects were presented with sentences, and had to write the target words and their meanings in blanks. The results of the productive measure showed that both higher and lower ability subjects who learned in context had greater mean percentage

scores (64.28, 37.2) than those who learned from word pairs (30.39, 28.5). However, the findings were only significant for the higher ability learners. Griffin attributes this to the way the information was processed. He suggests that learning in context may lead to more elaborate processing than decontextualized learning if learners have sufficient ability. This is supported by psychological theory on Elaboration (Craik and Lockhart, 1972; Craik and Tulving, 1975; Eysenck, 1978). The results of the receptive measure showed little difference between the two tasks with the higher ability subjects. However, the lower ability subjects who learned from word pairs (34.62) produced significantly larger gains than those who learned from sentences (22.1). Griffin attributes this to the similarity between the learning and test conditions, and the inability of less advanced learners to make use of context.

Laufer and Shmueli (1997) compared the effects of learning 20 vocabulary items from word pairs, glossed sentences, glossed texts, and glossed elaborated texts. They describe the word pair and sentence learning tasks as focus oriented, and the text and elaborated text tasks as quasi-incidental because the text groups did not know that the purpose of the task was vocabulary learning. The subjects were high school students studying English as a foreign language. The text groups were told to concentrate on comprehension while the other groups were told to learn the words. The text groups were given 55 minutes to complete their treatments, and the sentence and word pair groups were given 10 minutes. A surprise multiple-choice test measuring knowledge of meaning and form was used to determine vocabulary learning. They found that the sentence (17.06) and elaborated text (15.32) conditions resulted in significantly higher mean scores than the word pair (14.05) and text (12.85) groups immediately after the treatment.

certain whether the findings were the result of the treatments, the consolidation task or a combination of the two.

Seibert (1930) found that learning from word pairs was more effective than learning from sentences with translations of the target words given, and a combination of the two tasks. Productive knowledge of meaning and form was measured with two tests. The first test was a translation test in which learners were given the L1 translations and had to write the L2 target words. On the second test, learners were given the original sentence contexts with the L1 translations and again had to write the L2 target words. The word pair group had the highest mean scores on both tests and the sentence context group had the lowest mean scores. The tests were administered fifty minutes after the treatments. The word pair group recalled an average of 33.5 words out of a possible 48 on the translation test, and 33.4 on the cued recall test while the sentence context group had mean scores of 28.2 and 28.6 on the two tests. Unfortunately, the data was not analyzed to determine whether the findings were statistically significant.

Taken as a whole, the results of Dempster (1987), Griffin (1992), Laufer and Shmueli (1997), and Seibert (1930) indicate that both learning from word pairs and glossed sentences may be effective methods of learning vocabulary. Learners from both conditions produced large gains in knowledge of meaning and form in a relatively short time. The studies do not, however, show one task to be superior nor do they demonstrate that context has any effect on vocabulary learning. Dempster (1987), and Laufer and Shmueli's (1997) data indicates that context may have little or no effect on vocabulary gains. Griffin (1992) suggests that contextualized tasks may be superior for advanced learners, and inferior for beginners, and Seibert (1930) reports that decontextualized learning is superior. The contrasting results indicate that further research is needed to clarify the findings. One weakness of the research is that only knowledge of meaning and form was measured. It could be expected that tasks that provide the meaning of target words will produce large gains in knowledge of meaning. It may be more useful to determine how other aspects of knowledge are affected. Learning in context may contribute to gains in knowledge of syntax, collocation, and

grammatical functions that would appear to be ignored in decontextualized tasks. It may be necessary to measure all of these aspects to find how context affects vocabulary acquisition.

3.2 What may be gained from learning in context?

Past research has demonstrated that learners can gain knowledge of meaning and form through incidental reading tasks (see for example, Nagy, Herman, and Anderson, 1985), and intentional reading tasks (Gipe and Arnold, 1979). Unfortunately, very few studies have investigated whether any other aspects of knowledge are gained from meeting words in context. This is due to neglect rather than contrasting results. This is surprising considering the strong support from many vocabulary researchers for learning in context (see for example, Crow, 1986; Krashen, 1989; Oxford and Crookall, 1990).

Certainly, learners can gain knowledge of meaning from context. In fact they may learn much more about the meaning of a word from context than they can from a translation or synonym. Often it is context that makes the meaning of a polysemous word apparent. Miller (1999) reports that each meaning sense of a word may be associated with different contexts. One of his examples is the noun *shot*. In the sentence 'That was a good shot,' the meaning of *shot* is dependent on the discourse in which it is used. In a bar, *shot* refers to a drink, to a photographer it refers to an image, and to an athlete it is something entirely different. It is highly unlikely that in a decontextualized learning task, learners could gain such knowledge from a synonym or L1 translation.

The different semantic relationships and associations words have might also be learned through context. In a decontextualized task, those associations may be dependent on the associations that have developed with that word in the L1. Knowledge of the grammatical functions of a word such as its part of speech, derivatives, and the places it usually appears in a sentence, as well as knowledge of syntax such as syntagmatic

associations, and collocations may also be gained through context. Intuitively, syntactic and grammatical knowledge do not seem as likely to be gained through decontextualized learning, and one criticism of tasks such as learning through word pairs is that they do not provide learners with information on how to use target vocabulary (Oxford and Crookall, 1990). However, there is some evidence that decontextualized learning may also contribute to knowledge of grammatical functions, and syntax (see, Experiment 1; and Section 3.3 below). Learners may also learn how to use a word through meeting it in different contexts. Unless there is a large degree of overlap in L1 meaning, it seems unlikely that this would be gained through decontextualized learning.

In summary, it seems likely that many aspects of vocabulary knowledge may be gained through learning in context. Moreover, context may provide a better chance of gaining vocabulary knowledge than decontextualized learning from translations, definitions or synonyms. However, this is not what occurred in Experiment 1. Although it was found that incidental learning from reading led to significant gains in knowledge of orthography, meaning and form, association, syntax, and grammatical functions, when the time of the tasks was equal, decontextualized learning was found to be more effective in promoting knowledge of those aspects.

The following section briefly examines Experiment 1 and provides an explanation for why learning from a decontextualized task may be found to be superior to learning in context.

3.3 Experiment 1

Experiment 1 investigated incidental vs. explicit learning. The results showed that learning from word pairs may lead to large gains in knowledge of orthography, association, syntax, grammatical functions, and meaning and form in a relatively short time. While it might be expected that paired associate learning would lead to significant gains in orthography, meaning and form, and perhaps association (only

paradigmatic associations- words that had a semantic link to the target word – were included in this category), it is surprising that learners also produced large gains in knowledge of grammatical functions and syntax. In fact, gains in productive knowledge of grammatical functions and syntax made by the paired associate learners were greater than for incidental learners who met target words in context 10 times. This suggests that context may have relatively little effect on knowledge of grammatical functions, and syntax in comparison to knowing meaning and form. Since there were confounding variables it is impossible to determine whether this is correct. The present study seeks to shed some light on this.

One possible explanation for the findings was that the subjects used their knowledge of the target word's L2 synonyms as models and transferred their syntactic and grammatical knowledge from the models to the target words. Since the target words were nonsense words they could only use knowledge of the paired associates to answer questions. For example, the target nonsense word *masco* was paired with *train*. Since the subjects had no L2 knowledge of *masco*, they were likely to assume that it was a less frequent synonym of *train* such as locomotive. It seems unlikely that they would be able to use their L1 knowledge of *train* to correctly answer the productive knowledge of syntax, and grammatical functions tests because of differences between English and Japanese syntax and grammar. Therefore, they may have used their L2 grammatical and syntactic knowledge of *train* to respond on those tests. In the productive tests of grammatical functions the subjects were required to write each target word in a sentence with grammatical accuracy. The subjects may have created sentences containing the synonyms, and replaced the synonyms with the target words. For example, *masco* could replace *train* in the following sentences: 'The train leaves in 10 minutes,' 'The train is very slow,' and 'The train will get to the station at six o'clock.' Without L2 knowledge of the synonyms, the subjects may not have been able to demonstrate knowledge of syntax and grammar. If this is what occurred, it suggests that learners may gain knowledge of synonyms faster than non-synonyms. This is supported by Nation's (1990, 2001) theory of learning burden.

Nation argues that the amount of effort required to learn a word is different for different words, and for different learners. He refers to the effort needed to learn a word as its learning burden and states:

The general principle of learning burden is that the more a word represents patterns and knowledge that the learners are already familiar with, the lighter its learning burden. These patterns and knowledge can be from the first language, from knowledge of other languages, and from previous knowledge of the second language. So, if a word uses sounds that are in the first language, follows regular spelling patterns, is a loan word in the first language with roughly the same meaning, fits into roughly similar grammatical patterns as in the first language with similar collocations and constraints, then the learning burden will be very light (Nation, 2001: 23-24).

Learning burden provides a reasonable explanation for why the subjects in the first study produced large gains in knowledge of orthography, syntax, association, grammatical functions, and meaning and form despite learning from a task that focused solely on meaning and form. Synonyms may represent different aspects of knowledge that have already been acquired. Although the degree of overlap of knowledge varies from synonym to synonym, the fact that there is at least partial overlap may facilitate acquisition.

The learning from synonyms hypothesis contrasts with Laufer's (1990a) argument that synonymy is one of seven interlexical factors that can reduce the chances of vocabulary acquisition. Laufer gives two reasons why synonyms may be more difficult to learn than other words. First, learners often make mistakes using synonyms because some of them may be substituted effectively in some contexts but not in others. For example, 'strong' and 'powerful' have similar meanings but only tea can be 'strong' and engines may be 'powerful' but not 'strong'. This is an important point and demonstrates that synonyms may have different learning burdens. Certainly some words such as 'good' and 'nice' are synonymous in many contexts while others such as 'powerful' and

'strong' are synonymous to a lesser degree. Although they may not translate in every context they are synonymous in some and this may reduce learning burden and facilitate acquisition. The second reason she gives is that less advanced learners are unlikely to make an effort in learning words with similar meanings when they have a greater need to learn unknown L2 meanings. This is a valid point with important pedagogical implications but it relates more to learner motivation than word difficulty.

At present research has not addressed whether or not learners gain knowledge of words that have known synonyms faster than non-synonyms. Further research examining this issue would be valuable for three reasons. First, it would help vocabulary researchers better understand the vocabulary acquisition process. Current theories of vocabulary learning suggest that young L1 learners incidentally acquire words at a phenomenal rate through reading and listening (Nagy, Herman, and Anderson, 1985). However, there is little empirical support for this. If synonymy facilitates acquisition, it would help to support the incidental learning hypothesis. Furthermore, it would indicate that other types of words with lower learning burdens might be acquired more easily. Second, determining whether synonyms are acquired faster than other words would be very useful for teachers and learners. Teaching and learning strategies that strengthen links between synonyms could be used to improve learning. Third, it would demonstrate that learning burden is an important indicator of the rate of acquisition. If this is correct, it would greatly improve our understanding of word difficulty.

The present study investigated the effects of synonymy on vocabulary learning. Specifically it examined the following questions:

1. Do learners gain more knowledge of synonyms than non-synonyms from the same learning task?
2. How are aspects of vocabulary knowledge affected by synonymy?

3.4 Implementation of the research

3.4.1 Research Questions

The present study will address the following questions:

1. Is learning from word pairs as effective as learning from glossed sentences?

Experiment 1 showed that learning from word pairs might be more effective than learning in context. However, in that study the learners who met target words in context were not given the meanings of the target words. Providing the meanings in the contextualized task may show that learning from glossed sentences contributes to larger gains in some aspects of knowledge such as syntax and grammatical functions.

2. Which aspects of knowledge are gained through learning from word pairs and learning from glossed sentences?

Experiment 1 demonstrated that knowledge of orthography, syntax, grammatical functions, association, and meaning and form may be gained through incidental vocabulary learning and learning from word pairs. This experiment may reconfirm that knowledge of those aspects might be gained through learning from word pairs, and show that knowledge of those aspects is also gained through learning from glossed sentences.

3. How does the existence of a known L2 synonym affect vocabulary knowledge?

Laufer (1990a) suggests that synonymy hinders vocabulary learning. The results from Experiment 1 suggest the opposite. Experiment 2 may confirm that synonymy promotes vocabulary knowledge.

4. Are words acquired at different rates?

Nation's (1990, 2001) theory on learning burden suggests that different words are acquired at different rates. If synonymy is found to promote vocabulary learning that would indicate that synonyms are learned more easily than non-synonyms supporting Nation's argument.

5. How does context affect knowledge of orthography, meaning and form, association, syntax, and grammatical functions?

Previous research comparing decontextualized and contextualized vocabulary learning has shown that context has little effect on vocabulary gains (Dempster, 1987; Griffin, 1992; Laufer and Shmueli, 1997; Pickering, 1982; Prince, 1996; Seibert, 1930). However, in all of those studies only knowledge of meaning was measured. The results of Experiment 1 indicate that context may have relatively little effect on gaining knowledge of association, meaning, grammar, and syntax. However, in that experiment incidental learning from context was compared with explicit learning from word pairs. Experiment 2 may show that context does promote larger gains of some aspects such as syntax and grammatical functions.

3.4.2 Pilot Studies

Pilot studies conducted with Japanese learners studying English as a foreign language determined the following:

1. The subjects were unlikely to have prior L2 knowledge of the 20 target words. However, they were likely to know high frequency L2 synonyms of 10 of the target words.
2. The 10 nonsense words created for this experiment were unlikely to cause any interlanguage problems for the subjects.

3. The subjects should be able to understand the contexts and running words used in the experiment.
4. Eight minutes was enough time for the subjects to complete the treatments.
5. Japanese learners were able to correctly answer all of the receptive tests when they knew the L1 meanings for each target word.

3.4.3 Subjects

The population sample consisted of 84 volunteers learning English as a foreign language in Fukuoka, Japan. The subjects were selected from two first-year EFL classes at Kyushu University. All of the subjects had studied English for a minimum of six years, and had scored 80 percent or higher on the 2000 word level (a frequency band containing the 1001st to 2000th most frequent words) of Version 1 of the Vocabulary Levels Test (Schmitt, 2000). Their mean score was 27.1 indicating that they were well in control of that level, and that they had receptive knowledge of almost all of the 2000 most frequent words (Schmitt, Schmitt, and Clapham, 2001). The subjects were randomly assigned to the experimental and comparison groups.

3.4.4 Design

Two groups – experimental and comparison – were used in this study to examine the effects of context and synonymy on vocabulary learning. The experimental group participated in a learning from word pairs task, and the comparison group completed a learning from glossed sentences task.

The word pairs task involved learning a list of 20 nonsense words matched with their L1 translations. The subjects in the comparison group received the same list of word pairs; however, each target word was also shown in a sentence beside the word pairs. All of the target words were underlined in the sentences. Both groups were given eight minutes to learn the target words. After the conclusion of the treatments, a surprise

vocabulary test was administered. The same five knowledge types that were examined in the first study - orthography, association, meaning and form, syntax, and grammatical functions - were isolated and measured in a series of 10 tests. Nine of the ten tests were described in the first study. One test from the first study - receptive recognition of meaning and form - was replaced by a productive measure of meaning and form. The tests assessed each knowledge type at two levels of difficulty. Productive tests were designed to measure fuller knowledge, and receptive tests were designed to measure partial knowledge.

The subjects were not given any instruction on learning strategies such as the keyword technique prior to the treatment nor were they told how they would be tested. The testing procedure was the same as in Experiment 1. Each test was on one page. The first test measuring productive knowledge of orthography was administered to all of the subjects at one time because it involved aural cues spaced twelve seconds apart. The subjects were given as much time as they needed to complete the rest of the tests. When a subject finished a test it was collected and the next test was handed out.

3.4.5 Procedure

The procedure in this study was similar to the one used in the previous experiment. It was conducted in three stages over two weeks in two first year EFL classes at Kyushu University in Fukuoka, Japan. In the first stage of the experiment, potential subjects completed the Vocabulary Levels Test (Schmitt, 2000) to determine their L2 vocabulary size. In the second stage of the experiment, two subjects who did not take part in stage three were involved in case studies. The subjects were randomly assigned to the experimental group and the comparison group. They were asked to think aloud as they completed the treatments and the dependent measures to provide insight into their thought processes. Their results were not included in stage three. In the third stage of the experiment, the subjects completed the treatments and the dependent measures.

The experimental procedure was implemented over the following three stages:

- Stage 1 The Vocabulary Levels Test was administered to determine the vocabulary size of potential subjects. Subjects were randomly assigned to the treatment groups.
- Stage 2 The case studies were conducted.
- Stage 3 Subjects completed the treatments and dependent measures.

3.4.6 Stage 1 procedure

The Vocabulary Levels Test (Schmitt, 2000) was administered to potential subjects. At the outset, the test was explained and any questions regarding the test were answered to make certain that the instructions were completely understood. Subsequently, volunteers who had scored 80 percent or higher at the 2nd 1000 word level were randomly assigned to the two treatments. The Vocabulary Levels Test was used to ensure that all subjects in the study had a minimum vocabulary size. The Vocabulary Levels Test and pilot studies made it reasonably certain that the subjects would know all of the words presented in the texts to the comparison group.

3.4.7 Stage 2 procedure

Case studies were done with two volunteers. Neither of the subjects was involved in any of the classes that participated in Stage 3. The purpose of the case studies was to gain insight into the strategies that subjects employed to learn the target words, and the methods they used to complete the dependent measures. The subjects were randomly assigned to either the experimental or comparison groups. Each subject was asked to think aloud as they did the treatments and the dependent measures. The procedure in stage 2 was identical to stage 3 except that the subjects verbalized their thoughts. After the dependent measures were completed, subjects were asked a number of questions to clarify the findings. During stage 2 each subject was audio-recorded.

3.4.8 Stage 3 procedure

At the outset, instruction was provided on how to complete the treatments. The treatment was discussed in the subject's L1 to make certain that the instructions were completely understood. After the tasks were completed, the dependent measures were administered. When a test was completed it was collected, and then the next dependent measure was administered.

3.4.9 Materials

3.4.10 Target Words

Twenty words were chosen for the experiment. 10 were low frequency words with high frequency synonyms (Set A), and the other 10 were low frequency words without high frequency synonyms (Set B). L2 target words with high frequency synonyms allow subjects the possibility of gaining L2 knowledge from both their L1 translations and their L2 synonyms. In contrast, using target words without high frequency synonyms ensured that subjects could only gain knowledge of those words from their L1 translations or the contexts in which they were presented. Comparing the results for the two sets of words may determine the effects of synonymy. Both sets of target words were comprised of six nouns and four verbs. The ratio of six nouns to four verbs was used because nouns and verbs are the most common parts of speech found in natural text, and the 6:4 ratio approximates their proportional frequency of occurrence (Kucera and Francis, 1967). All of the target words were taken from the fifth frequency band in the COBUILD dictionary. It accounts for the 6601st to 14,700th most frequent words. The target words with high frequency synonyms were: locomotive, visage, lane, abode, boulder, crave, doze, sob, abhor, and dagger. The target words without high frequency synonyms were: lick, spear, recluse, pawn, landfill, mourn, convent, pier, reef, and marinate.

3.4.11 Nonsense words

Nonsense words replaced the L2 forms of the 20 target words. The 10 nonsense words from the first study were used and an additional 10 nonsense words were created. The same criteria from the first study were used to create the new words. All of the nonsense words were two syllables, and resembled English words phonetically and orthographically. In pilot tests, subjects were able to pronounce all of the words correctly, and indicated that none of the words resembled any L1 or L2 words phonetically. In each frequency group, seven of them were five letters long, and three were six letters long. The twenty nonsense words were: *ancon*, *cader*, *dangy*, *denent*, *faddam*, *hodet*, *masco*, *pacon*, *sagod*, *tasper*, *copac*, *gishom*, *hattaw*, *ictay*, *mesut*, *nasin*, *nuggy*, *tagon*, *toncop*, and *dapew*.

3.4.12 Word Pairs

In the word pairs treatment, the twenty nonsense words were presented together with their L1 translations on one page (see Appendix D). The nonsense words were on the left of the translations. In the following example, the target word *locomotive* was replaced with the nonsense word *masco*, and paired with the Japanese translation of locomotive:

| | |
|--------------|-----|
| masco | 機関車 |
|--------------|-----|

The order of the word pairs, pairs with high frequency synonyms and pairs without high frequency synonyms, was controlled to ensure that it did not affect learning. Target words with high frequency synonyms alternated with target words without high frequency synonyms. For half of the experimental subjects the high frequency nonsense words were presented first, and for the others the low frequency nonsense words were first.

3.4.13 Sentence Contexts

In the learning from glossed sentences treatment, the comparison group was presented with the 20 target words followed by their L1 meanings. One sentence containing a target word followed each word pair (see Appendix E). The following example is from the learning from the glossed sentences treatment:

masco 機関車 The driver got off the **masco**.

The sentences were selected from the British National Corpus. Three factors were considered when selecting the contexts: the number of words, the frequency of the words, and the ease in comprehending the sentence. Longer sentences were not chosen because they increased the time necessary to complete the treatment. The average length of the sentences for the words with high frequency synonyms was 9.2 words while the average for the words without high frequency synonyms was 12.1 words. Contexts that contained words that were unknown to any subject in the pilot tests were not used in the study. Controlling the running words in the contexts ensured that the learners would quickly understand the sentences. Extensive pilot testing was used to ascertain that subjects were likely to be familiar with all of the running words in the contexts. Since the purpose of the contexts was to determine how they influenced vocabulary knowledge, it was necessary that they were easily understood. If subjects could not understand some of the sentences, it would not be clear how context affected learning.

The order of the contexts was controlled in the treatment. Half of the subjects in the comparison groups met contexts containing target words with high frequency synonyms first, and half of the subjects met contexts without high frequency synonyms first. This was to ensure that their order was controlled for.

3.4.14 Controlling for interlexical and intralexical factors that may affect vocabulary learning

To ensure that any differences between the target words with high frequency synonyms and the target words without high frequency synonyms were due to synonymy, factors contributing to word difficulty were controlled. Many factors affect the ease with which words are learned (Higa, 1965; Lado, 1956; Laufer, 1990a; Ludwig, 1984, Tinkham, 1993; Waring, 1997a; Yoshikawa, 1998). In a series of studies, Laufer (1989, 1990a, 1990b, 1990c, 1997) identifies a large number of lexical factors that affect learning, and differentiates between them as being either interlexical or intralexical features. She defines interlexical factors as those that involve the relationships between a new word and L1 and L2 words already learnt, and intralexical factors as the intrinsic properties of a word. Laufer (1997) lists the following as intralexical features that may affect vocabulary learning: pronounceability, orthography, length, morphology, synformy, part of speech, abstractness, specificity/register restrictions, idiomaticity, and multiplicity of meaning. She suggests that word length, part of speech, and abstractness have no clear effect on acquisition. Laufer (1990a) lists the following as interlexical features that can affect vocabulary learning: degree of similarity to L1 in form and meaning, amount of overlap between the word in L1 and L2, degree of similarity in connotative meaning between the word in L1 and L2, semantic relationship with other words, and the degree of similarity in sound to other words in L2. Several steps were taken to control all of the intralexical factors and ensure that the only interlexical factor distinguishing the two sets of target words was synonymy.

The nonsense words were pilot tested to ensure that potential subjects could easily differentiate between them and any other L1 or L2 words phonetically and orthographically. The target words were selected and pilot tested to ensure that they all had a roughly equivalent degree of overlap in semantic grids between the L1 and L2, and had similar connotative meanings in the L1 and L2. It is difficult to ascertain the semantic relationships the target words may have with other words for individual learners because each learner is likely to have different paradigmatic associates for a

word. However, since all of the nonsense words represented very low frequency words in both L1 and L2, it is unlikely that the subjects had built up strong semantic relationships with any of the words.

Two forms of each treatment were prepared to ensure that nonsense words paired with one set of target words were not easier to learn than the other nonsense words. The first form of the treatments presented the original nonsense words paired with the first set of target words, and the new nonsense words paired with the second set of target words. In the second treatments, the pairings were reversed. Since the nonsense words used in the experiment were the same for both treatments, word pairs and sentence contexts, and both high and low frequency target words (using two forms of each treatment), pronounceability, orthography, morphology, and synformy could not affect any differences found between the variables. Specificity and register constraints are more likely to affect the later stages of acquisition and were therefore unlikely to be a factor in this study. To ensure that idiomaticity and multiple meanings did not influence vocabulary learning, none of the nonsense words met in sentence contexts involved idioms, and only one sense of a target word's L1 and L2 meaning was presented in the contexts.

3.4.15 Dependent Measures

Ten tests were used to measure the effects of synonymy, and context on vocabulary learning (see Appendix F). Nine of the tests were used in the first study. Only the receptive recognition test measuring knowledge of meaning and form was replaced by a productive test of meaning and form. Since the meaning and form of the target words were given to all of the subjects in this study, it was likely that a ceiling effect would occur if the recognition test had been used. The productive test was a much more demanding measure of meaning and form, and pilot tests indicated it would be a more useful test in this study. The tests in the second study measured the following:

Receptive and productive knowledge of orthographic form;
 Receptive and productive knowledge of meaning and form;
 Receptive and productive knowledge of grammatical functions;
 Receptive and productive knowledge of association;
 Receptive and productive knowledge of syntax.

As the first study indicated, it was important to have tests at different sensitivities to measure depth of knowledge. It was expected that the subjects would score higher on the receptive recognition tests. All of these tests used a multiple-choice format. The distractors were not too closely related, and allowed subjects to score correctly with partial knowledge of target words. The tests which measured productive knowledge, were much more demanding. None of these tests provided contextual clues to cue recall forcing the subjects to demonstrate a much fuller knowledge of the target words. Table 16 shows the tests that were used and the order in which they were completed. High and low frequency target words were always presented in different sequences on the tests to ensure that their order did not affect the results.

Table 16 Description and order of the dependent measures in Experiment 2

| Test order | Test type | Knowledge measured |
|------------|-----------------------------------|---|
| 1 | Spelling | productive knowledge of orthographic form |
| 2 | Multiple-choice | receptive knowledge of orthographic form |
| 3 | Translation | productive knowledge of meaning and form |
| 4 | Sentence construction | productive knowledge of grammatical functions |
| 5 | Write a syntagmatic associate | productive knowledge of syntax |
| 6 | Write a paradigmatic associate | productive knowledge of association |
| 7 | Multiple-choice | receptive knowledge of grammatical functions |
| 8 | Multiple-choice | receptive knowledge of syntax |
| 9 | Multiple-choice | receptive knowledge of association |
| 10 | Translation | receptive knowledge of meaning and form |

3.4.19 Productive and receptive measures of syntax

The tests that measured productive and receptive knowledge of syntax in the previous study were the same format in this study. The difference between the mean scores on the two measures indicated that the receptive test was more sensitive to partial gains in knowledge while the productive test measured fuller knowledge. Floor and ceiling effects were not found with any of the groups on the receptive test; however, a floor effect was found on the productive measure for the groups with fewer repetitions. This was not considered to be a weakness because the productive measure was designed to assess fuller knowledge. The scores for the word pair group suggest that both tests would be effective in this experiment.

3.4.20 Productive and receptive measures of grammatical functions

The same format used to measure productive and receptive knowledge of grammatical functions in the first study was also used in this study. The results of the first study were consistent with the measures of association, and syntax. The receptive test was sensitive to partial learning while the productive test measured fuller knowledge. The scores for the word pair group in the previous study suggest that there will not be a ceiling or floor effect in this study. Therefore, the tests are likely to be effective in this study.

3.5 Case studies

Case studies were done with two learners. One completed the word pairs task and the other completed the glossed sentences task. The subjects were asked to think aloud during the treatment and tests, and were interviewed after the tests were administered to clarify any misunderstandings.

The following observations emerged from the two case studies:

1. The subjects used different learning strategies in the two treatments. In the word pairs treatment, the subject used both a receptive word pairs task – covering the L1 meanings and then trying to recall them – and mnemonics to try to learn the target words. In the learning from glossed sentences treatment, the subject attempted to memorize both the target words and the sentences.
2. Both subjects indicated that their inability to recall which meanings were linked to each target word caused the most difficulty on the tests of meaning, association, syntax, and grammar.
3. The focus on spelling in the first two tests may have weakened the link between form and meaning. Both subjects reported that they forgot which words were linked to some of the meanings when they were concentrating on the tests of orthography.
4. Both subjects indicated that the set of target words with high frequency synonyms were easier to learn than the set without high frequency synonyms. However, individually this was not necessarily the case. When asked to select the easiest words, 7 of the first 10 words were synonyms. The subjects reported that their selections were based on synonymy, form, and interest.

3.6 Results

The descriptive statistics (means, standard deviations, and number of subjects) of vocabulary knowledge scores for the 10 dependent measures are shown in Table 17. To determine whether there were any overall differences among the treatment groups, a multivariate analysis of variance (MANOVA) was performed using the scores on the 10 dependent measures (productive orthography, receptive orthography, productive meaning and form, receptive meaning and form, productive association, receptive association, productive syntax, receptive syntax, productive grammar, and receptive

grammar). The independent variable was the type of learning task (learning from word pairs, and learning from glossed sentences). The MANOVA did not realize an overall statistically significant difference (Pillai's Trace, Hotelling's Trace, Wilks' Lambda, and Roy's Greatest Root) below $p < .05$ on any of the dependent measures ($F(10,73) = 0.83$, $p = .599$) indicating that context may have little effect on vocabulary. A summary of the statistical findings is shown in Table 18.

Table 17 Means and standard deviations of learning conditions on dependent measures in Experiment 2

| Learning Condition | <i>n</i> | PO | RO | PM | RM | PA | RA | PS | RS | PG | RG |
|--------------------|----------|-----------------|-----------------|-----------------|-----------------|----------------|-----------------|----------------|-----------------|----------------|-----------------|
| Word pairs | 43 | 10.98 (4.22) | 14.70 (3.64) | 11.21 (5.33) | 11.16 (5.56) | 8.35 (4.50) | 12.00 (4.78) | 8.12 (4.86) | 11.70 (4.95) | 9.79 (5.38) | 13.81 (4.77) |
| Contexts | 41 | 10.32 (4.53) | 14.83 (3.45) | 10.46 (5.13) | 11.83 (5.73) | 8.88 (5.07) | 12.93 (4.10) | 7.65 (4.65) | 13.34 (3.81) | 9.83 (5.39) | 14.80 (3.43) |

Note. Maximum score = 20, standard deviations are in parentheses, Key. PO=productive knowledge of orthography, RO=receptive knowledge of orthography, PM=productive knowledge of meaning and form, RM=receptive knowledge of meaning and form, PA=productive knowledge of association, RA=receptive knowledge of association, PS=productive knowledge of syntax, RS=receptive knowledge of syntax, PG=productive knowledge of grammar, RG=receptive knowledge of grammar.

Table 18 Summary of multivariate analysis of variance on vocabulary tests

| | productive | receptive |
|-------------|-------------------------------|-------------------------------|
| orthography | $F(1,82) = .477$, $p = .492$ | $F(1,82) = .029$, $p = .865$ |
| meaning | $F(1,82) = .426$, $p = .516$ | $F(1,82) = .292$, $p = .590$ |
| association | $F(1,82) = .001$, $p = .974$ | $F(1,82) = 1.18$, $p = .280$ |
| syntax | $F(1,82) = .194$, $p = .660$ | $F(1,82) = 2.89$, $p = .093$ |
| grammar | $F(1,82) = .257$, $p = .614$ | $F(1,82) = .905$, $p = .344$ |

Note. * $p < .05$, ** $p < .01$, *** $p < .000$

Table 17 shows that both tasks contributed to gains in receptive and productive knowledge of each aspect. Surprisingly there was very little difference between the size of the gains for the two tasks on each test. The larger gains on the receptive tests indicate that the receptive tests were more sensitive to gains in partial knowledge as was expected. The smallest mean percentage gains for both the word pairs (41%) and glossed sentence (38%) tasks on the productive tests were found on the measure of

syntax while the largest gains were found on the test of meaning (56%, 52%, respectively). The smallest gains on the receptive tests were found on the measure of meaning (56%, 59%) and the largest were found on the test of orthography (74%, 74%). It is likely that the translation format on the receptive test of meaning resulted in slightly lower scores than on the other receptive recognition tests. Recognition measures tend to be less demanding than recall measures because the format on multiple-choice tests allows subjects to score correctly through guessing (Nation, 2001).

Since none of the differences between the two tasks were found to be significant, it would appear that a single context has little affect on knowledge of syntax, and grammar. However, a comparison of the gains made from the tasks for target words with high frequency synonyms, and those without high frequency synonyms suggests that context may have a small affect on receptive knowledge of syntax. Table 19 shows the means, and standard deviations for synonyms and non-synonyms from the two treatments, and the combined treatments. The largest difference (1.06) between the two tasks is for the non-synonyms on the receptive test of syntax. The percentage change in receptive knowledge of syntax resulting from context was calculated to be 19 percent using Sackett's (1991) formula :

$$\% \text{ change} = d_1 \times 100 \times SD_{(\text{word pairs})} / M_{(\text{word pairs})}$$

Where:

$$d_1 = (M_{\text{glossed sentences}} - M_{\text{word pairs}}) / SD_{(\text{word pairs})}$$

A MANOVA was performed to determine whether there were any differences between the two tasks for synonyms and non-synonyms. Although a significant difference was found on the receptive test of syntax for words without high frequency synonyms ($F(1,82) = 4.20, p=.044$), the MANOVA did not realize an overall statistically significant difference (Pillai's Trace, Hotelling's Trace, Wilks' Lambda, and Roy's Greatest Root) below $p<.05$ between the two tasks ($F(20,63) = 1.03, p=.444$). Taken as a whole, the findings suggest that context may facilitate gains in knowledge of syntax.

However, coupled with the earlier result showing that there were no significant differences between the two tasks for any of the aspects, context may only have a slight effect on syntax which may be outweighed by other factors such as synonymy.

Table 19 Means and standard deviations of learning conditions on dependent measures for words with and without high frequency synonyms

| | | Word Pairs | | Context | | All Subjects (word pairs and context combined) | |
|----|--------------|------------|--------|---------|--------|---|--------|
| PO | Synonyms | 5.56 | (2.24) | 5.37 | (2.61) | 5.46 | (2.41) |
| | Non-synonyms | 5.42 | (2.27) | 4.95 | (2.52) | 5.19 | (2.39) |
| RO | Synonyms | 7.53 | (2.02) | 7.71 | (1.82) | 7.62 | (1.91) |
| | Non-synonyms | 7.16 | (2.16) | 7.12 | (2.01) | 7.14 | (2.08) |
| PM | Synonyms | 5.51 | (2.85) | 5.41 | (2.59) | 5.46 | (2.71) |
| | Non-synonyms | 5.70 | (2.73) | 5.05 | (2.86) | 5.38 | (2.80) |
| RM | Synonyms | 5.60 | (2.93) | 6.05 | (2.77) | 5.82 | (2.85) |
| | Non-synonyms | 5.56 | (2.81) | 5.78 | (3.11) | 5.67 | (2.94) |
| PA | Synonyms | 4.88 | (2.58) | 5.37 | (2.77) | 5.12 | (2.67) |
| | Non-synonyms | 3.47 | (2.21) | 3.51 | (2.66) | 3.49 | (2.42) |
| RA | Synonyms | 5.95 | (2.50) | 6.66 | (2.09) | 6.30 | (2.32) |
| | Non-synonyms | 6.05 | (2.54) | 6.27 | (2.31) | 6.15 | (2.42) |
| PS | Synonyms | 4.37 | (2.64) | 4.27 | (2.76) | 4.32 | (2.68) |
| | Non-synonyms | 3.74 | (2.51) | 3.39 | (2.31) | 3.57 | (2.41) |
| RS | Synonyms | 5.98 | (2.61) | 6.56 | (2.26) | 6.26 | (2.45) |
| | Non-synonyms | 5.72 | (2.58) | 6.78 | (2.13) | 6.24 | (2.41) |
| PG | Synonyms | 4.88 | (2.76) | 5.12 | (2.64) | 5.00 | (2.69) |
| | Non-synonyms | 4.91 | (2.78) | 4.71 | (3.04) | 4.81 | (2.90) |
| RG | Synonyms | 6.86 | (2.60) | 7.66 | (1.62) | 7.25 | (2.20) |
| | Non-synonyms | 6.95 | (2.49) | 7.15 | (2.07) | 7.05 | (2.28) |

Note. Maximum score = 10, standard deviations are in parentheses, Key. PO=productive knowledge of orthography, RO=receptive knowledge of orthography, PM=productive knowledge of meaning and form, RM=receptive knowledge of meaning and form, PA=productive knowledge of association, RA=receptive knowledge of association, PS=productive knowledge of syntax, RS=receptive knowledge of syntax, PG=productive knowledge of grammar, RG=receptive knowledge of grammar.

Three Repeated Measures MANOVA were performed to determine the effects of synonymy on vocabulary knowledge for each treatment group, and for both groups combined. The within subjects variable was synonymy (words with high frequency synonyms, and words without high frequency synonyms). The MANOVA realized an

overall statistically significant difference (Pillai's Trace, Hotelling's Trace, Wilks' Lambda, and Roy's Greatest Root) between the two factors (words with high frequency synonyms, and words without high frequency synonyms) for the word pairs task ($F(10,33)=5.20$, $p<.000$), the glossed sentences task ($F(10,31)=5.32$, $p<.000$), and for the two tasks combined ($F(10,74)=9.73$, $p<.000$).

Table 20 Summary of multivariate analysis of variance on vocabulary tests for word pairs, learning from glossed sentences, and for the combined tasks

| | Word Pairs | Context | All Subjects (word pairs and context combined) |
|----|------------------------------------|------------------------------------|---|
| PO | $F(1,42) = .334$, $p=.566$ | $F(1,40) = 1.23$, $p=.275$ | $F(1,83) = 1.55$, $p=.216$ |
| RO | $F(1,42) = 1.40$, $p=.243$ | $F(1,40) = 4.93$, $p=.032^*$ | $F(1,83) = 5.40$, $p=.023^*$ |
| PM | $F(1,42) = .556$, $p=.460$ | $F(1,40) = 1.60$, $p=.214$ | $F(1,83) = .189$, $p=.665$ |
| RM | $F(1,42) = .048$, $p=.828$ | $F(1,40) = 1.55$, $p=.220$ | $F(1,83) = 1.05$, $p=.308$ |
| PA | $F(1,42) = 30.17$, $p=.000^{***}$ | $F(1,40) = 37.29$, $p=.000^{***}$ | $F(1,83) = 67.30$, $p=.000^{***}$ |
| RA | $F(1,42) = .145$, $p=.705$ | $F(1,40) = 2.36$, $p=.132$ | $F(1,83) = .652$, $p=.422$ |
| PS | $F(1,42) = 5.93$, $p=.019^*$ | $F(1,40) = 7.33$, $p=.010^*$ | $F(1,83) = 13.35$, $p=.000^{***}$ |
| RS | $F(1,42) = 1.18$, $p=.284$ | $F(1,40) = .414$, $p=.524$ | $F(1,83) = .013$, $p=.908$ |
| PG | $F(1,42) = .013$, $p=.911$ | $F(1,40) = 2.07$, $p=.158$ | $F(1,83) = 1.17$, $p=.283$ |
| RG | $F(1,42) = .121$, $p=.730$ | $F(1,40) = 5.23$, $p=.028^*$ | $F(1,83) = 1.30$, $p=.257$ |

Note. $^*p<.05$, $^{**}p<.01$, $^{***}p<.000$, Key. PO=productive knowledge of orthography, RO=receptive knowledge of orthography, PM=productive knowledge of meaning and form, RM=receptive knowledge of meaning and form, PA=productive knowledge of association, RA=receptive knowledge of association, PS=productive knowledge of syntax, RS=receptive knowledge of syntax, PG=productive knowledge of grammar, RG=receptive knowledge of grammar.

Taken as a whole, the results of the Repeated Measures MANOVA indicate that synonymy may facilitate gains in vocabulary knowledge. For the word pairs task significant differences were found on the productive measure of association ($F(1,42) = 30.17$, $p<.000$), and on the productive test of syntax ($F(1,42) = 5.93$, $p<.05$). For the glossed sentences task significant differences were found on the productive tests of association ($F(1,40) = 37.29$, $p<.000$), and syntax, ($F(1,40) = 7.33$, $p<.05$), and the receptive tests of grammatical functions ($F(1,40) = 5.23$, $p<.05$), and orthography

($F(1,40) = 4.93, p < .05$). When subjects from both tasks were combined, significant differences were found on the productive tests of association ($F(1,83) = 67.30, p < .000$), and syntax ($F(1,83) = 13.35, p < .000$), and the receptive test of orthography ($F(1,83) = 5.40, p < .05$). However, it should also be noted that since there was little difference between gains for words with known synonyms and words without known synonyms on seven of the 10 tests for the combined tasks, and because the size of the differences that were significant were relatively small, the pedagogical significance of the results may also be relatively small. A summary of the findings on the Repeated Measures MANOVA is shown in Table 20.

3.7 Discussion

The results of the present study (Experiment 2) suggest that a single context may have little effect on gaining vocabulary knowledge thus supporting the findings of Dempster (1987), and Laufer and Shmueli (1997), and contrasting with those of Griffin (1992) and Seibert (1930). Experiment 2 expanded upon the earlier methodologies by measuring receptive and productive knowledge of orthography, meaning, association, syntax, and grammatical functions. All of the previous studies had defined vocabulary learning by a subject's ability to demonstrate knowledge of meaning and form. No significant differences were found between the tasks on all 10 dependent measures. While the similarity between the scores on the tests of meaning could be expected, it was somewhat surprising that the contextualized learning task did not promote significantly larger gains in syntax, and grammar. This may be due in part to the subjects' ability to make large gains in all aspects of knowledge through the word pairs task. The size of the gains in knowledge of association, grammar, and syntax suggests that learners are also able to use L1 knowledge to gain substantial L2 knowledge of an unknown word. This is particularly apparent if we consider the findings with the non-synonyms.

The experimental design – substituting nonsense words for low frequency words without high frequency synonyms - made it highly unlikely that the learners could draw

on any L2 knowledge in the word pairs task. Therefore, any gains could be attributed solely to the presence of the L1 meaning. Table 21 shows the mean percentage gains for non-synonyms made by the word pairs subjects. The large gains made in knowledge of association, grammar, and syntax indicates that there is a close relationship between those aspects and meaning and form. Gaining knowledge of meaning and form may facilitate gains in association, syntax, and grammar. This may be dependent upon the degree of overlap between L1 and L2 meaning and syntax, as well as the student's level of language proficiency. For intermediate learners simply linking L1 meaning with L2 form may be one of the fastest methods of gaining vocabulary knowledge. However, it seems less likely that beginners would be able to demonstrate much L2 knowledge of association, grammar, and syntax because they may lack knowledge of L2 grammatical functions as well as the vocabulary size to recognize and produce paradigmatic and syntagmatic associates. Further research investigating how meaning affects vocabulary knowledge with learners of different proficiencies would be a very useful follow-up to this study.

Table 21 Mean percentage gains with the non-synonyms on the 10 dependent measures for the learning from word pairs group

| | Dependent measures | | | | | | | | | |
|------------|--------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | PO | RO | PM | RM | PA | RA | PS | RS | PG | RG |
| Mean gains | 54% | 72% | 57% | 56% | 35% | 61% | 37% | 57% | 49% | 70% |

Key. PO=productive knowledge of orthography, RO=receptive knowledge of orthography, PM=productive knowledge of meaning and form, RM=receptive knowledge of meaning and form, PA=productive knowledge of association, RA=receptive knowledge of association, PS=productive knowledge of syntax, RS=receptive knowledge of syntax, PG=productive knowledge of grammar, RG=receptive knowledge of grammar.

Interestingly, when the gains made through the two tasks for non-synonyms were compared there is some evidence that context does facilitate learning. The results show that learning from glossed sentences promoted significantly larger gains in receptive

knowledge of syntax ($p < .05$) for non-synonyms. Since it was less likely that the subjects had L2 knowledge of syntax for non-synonyms, they may provide a more accurate assessment of the effects of context. The difference between the two tasks for words with high frequency synonyms may be smaller because learners may gain some L2 knowledge of syntax from known synonyms. This result also suggests that researchers must be very careful when selecting target words in vocabulary acquisition studies. The type of words chosen, and their L2 relationships may determine the size of gains. While the comparison between non-synonyms indicates that context may affect learning, it should also be noted that this does not reflect authentic vocabulary learning. It is only in the early stage of language learning that students do not know the synonyms for the majority of the words that they learn, and Griffin (1992) suggests that beginners may be unable to use context effectively. Therefore, it would seem that a single context would rarely affect vocabulary knowledge.

As mentioned earlier, one purpose of this study was to investigate how synonymy affects vocabulary knowledge. The results indicate that learners may gain significantly more knowledge of words with known synonyms than non-synonyms. In the learning from word pairs condition, subjects demonstrated significantly more productive knowledge of association ($p < .000$), and syntax ($p < .05$). In the context condition, scores were also higher on the productive knowledge of association ($p < .000$), and syntax ($p < .01$) tests, as well as the receptive measures of grammar ($p < .05$), and orthography ($p < .05$). Moreover, when both groups were combined, significantly higher gains were found on the productive tests of association ($p < .000$), and syntax ($p < .000$), and the receptive test of orthography ($p < .05$). It is not surprising that both groups had higher scores on the productive test of association since synonyms were one of the types of paradigmatic associates that were scored as correct. However, the larger gains on the tests of syntax, and grammar provide stronger evidence that learners acquire knowledge of synonyms more easily than non-synonyms. This was also supported by comments made by subjects in both the pretests and the case studies. When asked which group of words were easier to use in a sentence the majority of subjects selected the words with known synonyms. Similarly when asked to select the

easiest and most difficult words to learn they usually chose synonyms for the easiest, and non-synonyms for the most difficult.

Taken as a whole, the results support the hypothesis that synonymy facilitates vocabulary learning and support Nation's (1990, 2001) theory of learning burden. However, this does not necessarily discount Laufer's (1990a) argument that synonyms are more difficult to learn than non-synonyms. The reason for this is that the present study investigated initial gains in vocabulary knowledge while Laufer was referring to gaining full knowledge of a synonym. Both theories could prove correct. Since synonyms have a lower learning burden than non-synonyms, students may instantly gain partial knowledge of association, grammar, and syntax for a word from its synonyms. This is an important gain in knowledge because it may provide learners with enough knowledge to use the word. However, this development may soon be impeded if errors occur when the word is used. Since few synonyms can be substituted in every context, it is likely that at some point a newly acquired synonym will be used incorrectly. At this time, learners may be discouraged from using synonyms, and gains in knowledge could be delayed to the point that acquisition of non-synonyms could progress more quickly. However, it is also likely that acquisition of synonyms may progress at a normal pace with learners gaining receptive knowledge through meetings in context, and productive knowledge through trial and error. In fact the same argument can be used with paired associate learning. The results of the present study suggest that learners may also make large initial gains in association, syntax, and grammar by linking L2 form to L1 meaning. The degree of overlap between L1 and L2 meaning, and syntax may determine how easily the word is learned. Initial gains in knowledge followed by a drop in performance, and then a return to targetlike use may be quite common in L2 learning, and is what Kellerman (1985) refers to as 'U-shaped behavior.'

The learning from synonyms result is valuable for three reasons. First, it provides evidence that a word's learning burden may determine how quickly it is acquired. This is useful for teachers and learners because if they know the likely rate of acquisition for

different words they can determine how much time they need to spend on an item. Second, the result provides some evidence in support of Nagy, Herman, and Anderson's (1985) incidental learning hypothesis. Determining that some words such as words with known synonyms may be acquired more quickly may help to explain why young L1 learners are able to acquire words at such a phenomenal rate. Third, it suggests that teaching and learning strategies that strengthen the links between synonyms may improve learning.

Another purpose of this study was to determine which learning task was most effective. The results of Experiment 2 suggest that both learning from word pairs and learning from glossed sentences are effective methods of learning vocabulary. Both tasks promoted relatively large gains in knowledge of orthography, association, meaning, syntax, and grammar in a short time. One reason that learning from word pairs may be as effective as learning from glossed sentences is that it focuses learners on establishing a link between L1 meaning and L2 form. In tasks involving context, learners may process meaning but their attention may not be directed towards creating a link between meaning and form (Pressley, Levin, and McDaniel, 1987). The results of this experiment indicate that by gaining knowledge of meaning and form learners may also gain knowledge of association, syntax, and grammar. Thus, the decontextualized task may be just as effective. However, since Griffin (1992) found that lower and higher level learners may respond differently to contextualized and decontextualized learning tasks, a replication of this experiment with beginners would be a useful follow up to this study.

3.8 Summary of main findings

The results of this study suggest:

1. A single sentence context has little effect on gaining vocabulary knowledge. No significant differences were found between the contextualized and decontextualized tasks on the 10 dependent

measures. However, a comparison of the tasks with non-synonyms did show that the learning from glossed sentences task contributed to significantly larger gains in receptive knowledge of syntax. This indicates that context can have an effect but it may depend on the learning burden for each word.

2. Intermediate learners are able to make substantial gains in knowledge of syntax, association, and grammar when they gain L1 knowledge of meaning.
3. It may be easier to learn synonyms than non-synonyms. The results showed that gains in knowledge of association, syntax, grammar, and orthography were larger for synonyms than non-synonyms.
4. Both learning from word pairs and learning from glossed sentences may be effective methods of learning vocabulary. Both tasks may promote gains in knowledge of orthography, syntax, meaning, association, and grammar.

3.9 Rationale for the next experiment

In Experiments 1 and 2, the learning from word pairs task has been compared with incidental learning from reading, and learning from glossed sentences. In both experiments, it has proven to be very effective. One weakness of the two experiments was that learners might have used different learning strategies in the word pairs task. Word pairs can be learnt receptively and productively. Research investigating learning from word pairs indicates that the type of learning, receptive and productive, affects the type and amount of knowledge gained (Griffin and Harley, 1996; Stoddard, 1929; Waring, 1997b). Learning word pairs receptively promotes larger gains in receptive knowledge of meaning while learning productively leads to larger gains in productive knowledge of meaning. How other aspects of knowledge are affected by receptive and productive learning has not been determined.

The aim of the following three experiments is to investigate how receptive and productive learning tasks affect vocabulary knowledge. Four different tasks will be examined: receptive learning from word pairs, productive learning from word pairs, writing a sentence, and reading three glossed sentences. The research would be useful for three reasons. First, since language learning is often skill oriented, determining which type of learning is best suited to improve reading, listening, speaking, and writing would be of great benefit for teachers and learners. Second, it may help researchers to better understand the relationships between the different aspects of knowledge. Since only knowledge of meaning was measured in other researchers' word pairs studies, it may be useful to find how the other aspects are affected. Third, it would be of use to teachers and students to determine whether different receptive and productive tasks affect learning in the same way. At present, the only studies comparing the effects of receptive and productive learning on vocabulary knowledge involved the learning from word pairs task.

CHAPTER 4

4.0 Introduction

The majority of vocabulary is learned receptively through reading or listening (Nagy, Herman, and Anderson, 1985; and see Experiment 1). When vocabulary is taught in the classroom, learning is also likely to be receptive. Teachers may tell learners the meaning of a word, provide a definition, or use the word in a sentence; however, they are less likely to have students use an item apart from spelling or pronouncing it. Vocabulary learning tasks are also more likely to be receptive than productive. Receptive activities such as looking up words in a dictionary, matching words with their meanings or definitions, guessing from context, and learning from word pairs (most often words are learned receptively in L2-L1 pairs) are more common than productive activities such as cloze exercises, and writing tasks. Receptive tasks may be more popular because they are easier to design, grade, and complete than productive tasks. However, it has never been demonstrated that receptive learning is more effective than productive learning; in fact research indicates that the opposite may be true.

The previous experiments in this thesis have shown that the learning from word pairs task may be very effective. Experiment 1 demonstrated that when time spent on tasks was equal, learning from word pairs contributed to greater gains in receptive and productive knowledge of orthography, meaning and form, association, syntax, and grammatical functions than incidental learning from reading. Experiment 2 showed that learning from word pairs was as effective as learning from glossed sentences. However, one weakness of the experiments is that they did not differentiate between receptive and productive learning from word pairs.

Research on learning from word pairs suggests that the type of learning, receptive and productive, affects the type and amount of knowledge gained (Griffin and Harley,

1996; Stoddard, 1929; Waring, 1997b). If words are learned receptively, then learners are likely to gain significantly more receptive knowledge while productive learning leads to larger gains in productive knowledge. This provides a possible explanation for why a learner's receptive vocabulary may be larger than his or her productive vocabulary which is what current research suggests (see, for example, Laufer, 1998; Laufer and Paribakht, 1998; Waring, 1997c). Since vocabulary learning is predominantly receptive, learners are more likely to gain receptive knowledge than productive knowledge.

To my knowledge, apart from the studies investigating learning from word pairs there has been no research comparing receptive and productive learning. This is surprising considering the implications of the word pair studies. If receptive learning is better suited to developing receptive knowledge, then it may not be efficient to use productive tasks in a learning program in which the primary aim is to improve receptive skills. Similarly if the goal is to increase production, receptive tasks may be less beneficial. However, if only one task is used, productive tasks may be more effective.

One weakness of the word pair studies (Griffin and Harley, 1996; Stoddard, 1929; Waring, 1997b) is that vocabulary gains were only measured with tests of meaning and form. Nation's (1990, 2001) vocabulary knowledge framework shows that there is much more to knowing a word than simply knowing its meaning. Further research examining which aspects of knowledge are gained from receptive and productive learning tasks is needed.

The present study seeks to investigate the effectiveness of receptive and productive learning tasks. Specifically it examines the following questions:

1. Which aspects of knowledge are gained through receptive and productive learning?
2. Which is more effective, receptive or productive learning?

4.1 What is receptive and productive vocabulary knowledge?

Despite numerous studies investigating receptive and productive vocabulary size, and several papers attempting to clarify the notions of reception and production, there are no clear and concise definitions of the concepts. What it means to know a word receptively and productively has varied from study to study, and these terms are often defined only by the tests used to measure them. There is not even agreement on the labels receptive and productive. They have also been referred to as passive and active, comprehension and production, and understanding and use. Most often receptive knowledge refers to understanding a word during reading or listening while production refers to using words in speech or writing. While tests measure reception and production as the product of learning they can also refer to the processes, skills, and abilities that are involved in understanding and using vocabulary (Waring, 1999). Nation (2001) points out that the terms may be misleading because receptive learning involves productive skills. For example, learning vocabulary through reading or hearing involves producing meaning. This difference between what is involved in learning a word, and the learning outcome is perhaps the largest problem with defining the terms. The lack of a clear definition has often led to a disparity between the construct definition and the test format (Melka Teichroew, 1982; Waring, 1999). The most common method of measuring receptive vocabulary has been through multiple-choice tests, self report tests, and matching tests while the standard methods of assessing productive vocabulary have been through cued recall tests, essay writing, and translation tests.

Almost all test formats can be criticized for two reasons. First, they involve both receptive and productive knowledge to some extent (Crow, 1986; Waring, 1999). In receptive tests such as multiple choice, translation, and matching tests, subjects need to produce the meanings of target words or distractors to recognize the correct response while in productive tests such as cued recall, and translation tests subjects must recognize the prompt to recall the target word. In consequence, there are many different testing methods, none of which appears ideal. Melka (1997) states that "it is

not obvious that any particular form of test is either specifically or adequately suited for testing either reception or production" (p.97). It is doubtful that any test will be without criticism until there are clear and concise definitions of reception and production. One solution would be to alter the definitions to conform to what is being tested.

The second criticism of the tests is that they tend to ignore most aspects of knowledge aside from meaning and form (Melka, 1997). There are numerous studies that have investigated receptive and productive knowledge of meaning and form (Griffin and Harley, 1996; Laufer, 1998; Laufer and Paribakht, 1998; Morgan and Oberdeck, 1930; Stoddard, 1929, Waring, 1997b; Waring, 1997c) while there appears to be only two (Schmitt, 1998; Schmitt and Meara, 1997) that look at both receptive and productive knowledge of other aspects of knowledge. The first two experiments in this thesis have shown that learners may gain knowledge of orthography, association, syntax, grammatical functions, and meaning and form from a vocabulary learning task. The present study seeks to determine the effects of receptive and productive learning on each of these knowledge types.

Another weakness of receptive and productive tests is that some involve more than one aspect of knowledge while others do not. This may make comparisons between tests or studies misleading. For example, any productive test that involves context such as The Lexical Frequency Profile (Laufer and Nation, 1995) or the Productive Vocabulary Levels Test (Laufer and Nation, 1999) may require subjects to demonstrate knowledge of meaning and form, orthography, grammatical functions and collocations while a decontextualized test such as a translation test requires only knowledge of meaning and form, and orthography. The first two experiments reported in this thesis have shown that this can be overcome by using multiple tests that isolate each aspect of vocabulary knowledge, and together provide a more complete assessment of how well a word is known.

In the present study, 10 tests were constructed to isolate and measure five different aspects of receptive and productive knowledge. The aspects of knowledge were form,

not obvious that any particular form of test is either specifically or adequately suited for testing either reception or production" (p.97). It is doubtful that any test will be without criticism until there are clear and concise definitions of reception and production. One solution would be to alter the definitions to conform to what is being tested.

The second criticism of the tests is that they tend to ignore most aspects of knowledge aside from meaning and form (Melka, 1997). There are numerous studies that have investigated receptive and productive knowledge of meaning and form (Griffin and Harley, 1996; Laufer, 1998; Laufer and Paribakht, 1998; Morgan and Oberdeck, 1930; Stoddard, 1929, Waring, 1997b; Waring, 1997c) while there appears to be only two (Schmitt, 1998; Schmitt and Meara, 1997) that look at both receptive and productive knowledge of other aspects of knowledge. The first two experiments in this thesis have shown that learners may gain knowledge of orthography, association, syntax, grammatical functions, and meaning and form from a vocabulary learning task. The present study seeks to determine the effects of receptive and productive learning on each of these knowledge types.

Another weakness of receptive and productive tests is that some involve more than one aspect of knowledge while others do not. This may make comparisons between tests or studies misleading. For example, any productive test that involves context such as The Lexical Frequency Profile (Laufer and Nation, 1995) or the Productive Vocabulary Levels Test (Laufer and Nation, 1999) may require subjects to demonstrate knowledge of meaning and form, orthography, grammatical functions and collocations while a decontextualized test such as a translation test requires only knowledge of meaning and form, and orthography. The first two experiments reported in this thesis have shown that this can be overcome by using multiple tests that isolate each aspect of vocabulary knowledge, and together provide a more complete assessment of how well a word is known.

In the present study, 10 tests were constructed to isolate and measure five different aspects of receptive and productive knowledge. The aspects of knowledge were form,

association, syntax, grammatical functions, and meaning and form. Isolating the different kinds of knowledge allowed a more accurate representation of what was gained through learning. In previous research, test formats often involve more than one knowledge type making it difficult to determine why learners were unable to answer a question. For example, in a sentence production task the testers are often attempting to find whether the learner being tested has gained knowledge of meaning and form. However, if the testee is unable to answer the question it may be because he or she is lacking grammatical, collocational, or orthographic knowledge.

Table 22 Definitions of receptive and productive knowledge

| Knowledge type | R/P | Definition of receptive and productive knowledge for each knowledge type |
|-----------------------|-----|--|
| orthography | r | The learner can recognize the correct spelling. |
| orthography | p | The learner can produce the correct spelling. |
| syntax | r | The learner can recognize syntagmatic associates. |
| syntax | p | The learner can produce syntagmatic associates. |
| grammatical functions | r | The learner can recognize if the item is being used with grammatical accuracy. |
| grammatical functions | p | The learner can use the word in a sentence with grammatical accuracy. |
| association | r | The learner can recognize paradigmatic associates. |
| association | p | The learner can produce paradigmatic associates. |
| meaning and form | r | The learner can recognize the correct L1 meaning. |
| meaning and form | p | The learner can produce the form of the word when given the L1 meaning. |

In this study, receptive and productive knowledge are defined by the learning outcome demonstrated by the tests. Since the completion of most tests involved receptive and productive processes, skills, and abilities, reception and production may perhaps be best defined as a product of learning. The definitions are shown in Table 22.

4.2 Do learners gain receptive knowledge before productive knowledge?

The majority of research investigating reception and production has focused on vocabulary size. Despite all of the uncertainties surrounding the two concepts, the research has led to two widely accepted assumptions:

1. Receptive knowledge precedes productive knowledge (Aitchison, 1994; Channell, 1988; Melka, 1997).
2. A person's receptive vocabulary is larger than his or her productive vocabulary (Aitchison, 1994; Crow, 1986; Laufer, 1998; Laufer and Paribakht, 1998).

The first assumption has led to contrasting theories that reception and production form a continuum (Melka, 1997), are independent of each other (Meara, 1990), or are subject to a threshold effect (Meara, 1996). While the assumption that reception precedes production seems logical, it is certainly misleading. Learners may often gain productive knowledge of some aspects of knowledge before they gain some types of receptive knowledge. Advanced learners are likely to gain productive knowledge of form and partial productive knowledge of grammatical functions from seeing or hearing the form of an unknown word. This may occur before they gain receptive knowledge of meaning. Incidental vocabulary learning research has shown that learners gain very little receptive knowledge of meaning from one meeting in context, and the results of Experiment 1 suggest that they are unlikely to gain any knowledge of meaning from one presentation in an uninformative context. For example, from the following sentence 'Their request for extradition was denied.' learners who have never met the word *extradition* are likely to gain little receptive knowledge of meaning but may be able write or pronounce *extradition* correctly based on their knowledge of the rules of spelling, and are also likely to learn that it is a noun from its suffix.

Most people have used words without fully understanding them. While the outcome may have been funny or embarrassing they may still have been able to demonstrate

types of productive knowledge such as form, collocation, and grammatical functions. Many L2 teachers will be able to give examples of learners demonstrating productive knowledge of meaning before they have gained receptive knowledge of grammatical functions. Schmitt (2000) provides another example, noting that because he had heard the word 'indict' several times he was able to understand and use it in conversation long before he was able to recognize it in print because its spelling was different from what he had imagined. This shows the importance of the learning context. How a word is learned may determine the type of knowledge gained.

The arguments given above suggest that some aspects of productive knowledge may be gained before some aspects of receptive knowledge. However, since knowing a word has traditionally been defined by meaning, it may be more accurate to restate the first assumption as 'receptive knowledge of meaning precedes productive knowledge of meaning.' Melka (1997) reports that intuition, research with children learning their L1, as well as studies estimating vocabulary size support this assumption. The following section examines four of the studies that have investigated receptive and productive vocabulary size.

4.3 Receptive and productive vocabulary size

Research investigating vocabulary size has consistently found that learners have a much larger receptive vocabulary than a productive one. Although design flaws have plagued the studies, the research gives insight into some of the more common methods of measuring receptive and productive knowledge, as well as the potential pitfalls of the test types.

One of the earliest studies was by Morgan and Oberdeck (1930). They measured the L2 vocabulary size of five classes of university students learning German. They found that the size of receptive vocabulary exceeded that of productive vocabulary at all five levels of difficulty. The results suggest that at lower levels receptive knowledge increases faster than productive knowledge and at later levels production develops

faster than reception but never to the point that it equals reception. Although the study is important because it is the first to examine the gap between receptive and productive vocabulary, two problems reduce the validity of the findings. First, the receptive test provided the possibility of scoring correctly through guessing while the productive test did not. The receptive measure was a five item multiple choice test while the productive test was a translation test. Therefore, if subjects had no receptive knowledge of a target item, they still had a 20 percent chance of scoring correctly on the multiple choice test while the translation test offered virtually no chance of guessing correctly. A second problem is that responses on the productive test were sometimes marked as incorrect despite the fact that subjects may have had productive knowledge of those items. For example, they report that the response *frohlich* was a correct translation of the cue *happy*; however, since it did not correspond with the German cue on the receptive test it was marked as incorrect. Since there was not a follow up interview, it could not be determined whether or not the subject could have responded with the target item. In summary, since both problems were biased in favor of reception, it is not clear whether receptive vocabulary exceeds productive vocabulary or if it is the result of the experimental design.

Three recent studies by Laufer (1998), Laufer and Paribakht (1998), and Waring (1997c) found that receptive vocabulary was larger than productive vocabulary supporting the earlier findings of Morgan and Oberdeck (1930). All three studies measured receptive and productive vocabulary size using the same tests and similar procedures. Receptive knowledge was measured using the Vocabulary Levels Test (Nation, 1990) while productive knowledge was measured using the Productive Vocabulary Levels Test (Laufer and Nation, 1999). A third test, The Lexical Frequency Profile (Laufer and Nation, 1995) was also used to measure 'free' productive knowledge in Laufer (1998) and Laufer and Paribakht (1998). Once again, however, the test formats were biased in favor of receptive knowledge bringing the results into question.

The Vocabulary Levels Test is a matching test in which subjects must select the correct definition or synonym for three words from one of six options. For example,

- | | |
|------------|-------------------------------------|
| 1. arrange | |
| 2. develop | _____ grow |
| 3. lean | _____ put in order |
| 4. owe | _____ like more than something else |
| 5. prefer | |
| 6. seize | |

The productive version of the Levels Test is a cued recall test that involves subjects completing a word in a context. To limit the answers to the target vocabulary, the first letters of the words are provided. For example,

There are a doz_____ eggs in the basket.
Every working person must pay income t_____.
The pirates buried the trea_____ on a desert island.

Scoring a correct response depends on whether the subject is able to recall the target word, and then spell it correctly. However, this is not as simple as it sounds. Waring (1997c) and Laufer (1998) score responses differently in respect to the grammatical form of the word (tense), and the orthographic form. The tests measure vocabulary knowledge at four different word frequency levels (2,000, 3000, 5000, and 10,000) as well as the University Word List (Xue and Nation, 1984) allowing a direct comparison between receptive and productive knowledge at each level.

At first glance the tests appear ideally suited to assessing the relationship between receptive and productive knowledge. However, there are three reasons why such a comparison may prove to be misleading. First, the matching test allows the learner a chance to score correctly without any knowledge. There is a 17% chance that a subject will guess correctly on one response with no knowledge of any of the six choices, and with knowledge of some of the words the possibility of guessing correctly increases. The productive version offers little or no chance to guess correctly. If a subject does not know the correct response based on the context, the only possibility of guessing correctly is by producing a response based on the spelling of the first letters of

the word which is very unlikely. Therefore, it is likely that learners will score higher on the matching test and demonstrate greater receptive knowledge than on the productive test of cued recall. Second, the receptive version of the Vocabulary Levels Test involves knowledge of form and meaning while the productive version may also depend on knowledge of grammatical functions (depending on the method of scoring). Thus, the productive version is more demanding because it requires subjects to demonstrate more aspects of knowledge than the receptive version. A third reason a comparison may be misleading is that the productive version of the Levels Test may actually test receptive knowledge. Morton (1979) reports that subjects are able to recognize words when one phoneme is inaudible. Furthermore, when a partially pronounced word was produced in context, subjects were often unaware that it was not produced completely. This leads Melka (1997) to state "it is, then, clear that the presence of partial information is often sufficient to recognize a word" (p.87). Morton's findings suggest that a test of cued recall in which the target word is partially presented could actually be a recognition test in which only productive knowledge of orthography is demonstrated. Since the receptive test provides a greater possibility of answering correctly without knowledge, and productive knowledge of orthography is not tested, the productive test could thus be considered a stricter test of receptive knowledge.

The preceding comments indicate that comparing the results of the two tests is likely to be misleading. However, they do not reduce the value of the two versions of the Vocabulary Levels Test. The tests are extremely useful for determining the size of a learner's receptive vocabulary, and indicating which types of words students need to learn.

In the test of 'free' productive knowledge involving the Lexical Frequency Profile, subjects were asked to write a short passage of 300 to 400 words about a given topic, which was then analyzed in terms of the words lexical frequencies. For example, if the passage contained 50 word families consisting of 30 from the first 1,000 most frequent words, 10 from the second thousand, five from the third thousand, and five from the

University Word List, the numbers could then be converted into the percentages to reflect that learners productive knowledge at the different word frequencies. In this case that would be 60% - 20% - 10% - 10%. The Lexical Frequency Profile is thus able to provide useful data about a learner's productive knowledge at different word frequencies but since it does not allow learners to demonstrate productive knowledge of specific items, it cannot be used in a comparison with tests that do.

The writing task used in the Lexical Frequency Profile is a purely productive task. However, scoring word knowledge in such a task is extremely difficult. Producing a word in context involves not only knowledge of orthography and meaning, but also syntax, collocation, and grammatical functions. Even other aspects of knowledge such as register, associations, and concept and referents could pose problems with scoring a word as known or unknown. For example, if the meaning of a word is demonstrated in the context, and it is used correctly grammatically but the register is inappropriate how should it be scored? Laufer and Paribakht (1998) scored some words as correct despite grammatical errors such as writing the wrong verb tense, or the incorrect plural or singular form. Small spelling mistakes, and writing words with the wrong prepositions were also marked as correct despite indicating limited knowledge of orthography and syntax. This highlights one of the biggest problems with receptive and productive vocabulary tests; receptive tests tend to focus on form and meaning while productive tests may involve more aspects of knowledge. It is necessary to either ignore the other knowledge types on both receptive and productive tests or test multiple aspects receptively and productively. As can be seen from the Lexical Frequency Profile, it is probably impossible to measure the different aspects of knowledge in one productive test. Therefore, several tests are needed.

In summary, although all of the studies investigating vocabulary size have found receptive vocabulary to exceed productive vocabulary, design flaws in the studies suggest that their results may be biased in favor of receptive knowledge. The principal flaw has been that the receptive and productive tests have not been equally sensitive to gains in knowledge. While they may effectively demonstrate that a learner has gained

receptive or productive knowledge, unless they are equally sensitive an accurate comparison cannot be made. Further research is needed to determine whether: (a) receptive vocabulary is indeed larger than productive vocabulary, and (b) reception precedes production.

At present, there do not appear to be many test formats that are suited to testing both reception and production. In a review of receptive and productive testing, Waring (1999) reached a similar conclusion, and developed a self-report measure called the State Ratings Task. While it offers the potential to determine receptive and productive vocabulary size, it would not be appropriate for this study because it does not provide verifiable data on individual aspects of knowledge. Translation tests may be the best measure of meaning and form because the same format can be used. Since other aspects of knowledge have rarely been tested receptively and productively, new testing methods may need to be developed. This would be a useful area for future research. However, as is the case with meaning and form, it may prove to be extremely difficult to find receptive and productive measures that are equally sensitive. Since the present study seeks to compare receptive and productive learning rather than the size of receptive and productive gains, equally sensitive tests are not necessary. Comparing the mean scores on the receptive measures, and the mean scores on the productive measures may show which learning task is more effective.

4.4 Learning from word pairs research

Most research investigating reception and production has focused on either receptive and productive vocabulary size or whether receptive knowledge is gained before productive knowledge. Surprisingly, there is very little research that compares receptive and productive learning. However, research on learning from word pairs sheds some light on this issue.

Stoddard (1929), Griffin and Harley (1996), and Waring (1997b) compared learning vocabulary receptively in L2-L1 word pairs, and productively in L1-L2 pairs. In all

three studies, learning a word receptively involved meeting a decontextualized L2 vocabulary item, and then trying to recall its L1 meaning while productive learning involved meeting a decontextualized L1 item, and then trying to recall its L2 meaning and form. In the Stoddard, and Griffin and Harley studies, the target words and meanings were presented in word pairs, and in the Waring study they were presented on word cards. Gains in knowledge were measured with translation tests that mirrored the learning tasks. In the receptive tests, the subjects were given the L2 forms of the target words and had to supply their L1 meanings. In the productive tests, the subjects were given the L1 meanings and had to write their L2 forms.

In Stoddard's (1929) study, 328 junior high school and high school subjects learned 50 foreign language target items either receptively (L2-L1) or productively (L1-L2). The receptive test was a 25 question L2-L1 translation test and the productive test was a 25 item L1-L2 translation test. Stoddard found that subjects scored significantly higher on the receptive test (14.1) than on the productive test (7.0). This led Stoddard to suggest that receptive knowledge is easier to gain than productive knowledge. However, he also found that the subjects who learned productively (8.0) significantly outperformed the receptive subjects (6.0) on the productive measure. Similarly, the receptive group did significantly better on the receptive test (15.1) than the productive group (13.1). Overall, there was little difference between the scores on the two tests for the subjects who learned productively (20.9) and those who learned receptively (21.1). The results indicate that the method of learning, receptive or productive, may determine the type and amount of knowledge gained. Receptive learning may lead to larger gains in receptive knowledge than productive learning, and productive learning may result in larger gains of productive knowledge than receptive learning.

Griffin and Harley (1996) had four groups of higher and lower ability ESL learners study 20 word pairs productively and another four groups study the word pairs receptively for eight minutes. The target words were items which were believed to be unknown to the subjects. Two receptive and two productive groups completed the productive translation test, while the other four groups completed the receptive

translation test. Since each group took only one test, the possibility of a learning effect on a second test was eliminated. Griffin and Harley found that there was little difference between the combined mean scores on both measures for receptive (48.2, 29.55) and productive learning (45.19, 28.52), and suggest that they are equally effective tasks. The results also showed that gains in knowledge were highest if the learning task matched the testing type. Mean scores on the productive test were 52.98 and 30.66 for the productive groups and mean scores on the receptive test were 66.41 and 48.97 for the receptive groups. When the learning task did not match the test, the subjects who learned productively had significantly higher scores on the receptive measure (37.41, 26.39) than the subjects who learned receptively but were tested productively (30.0, 10.14). This indicates that if only one type of learning task is used, a productive task may be more beneficial. Griffin and Harley also found that the mean scores on the receptive tests (51.91, 37.68) were significantly higher than on the productive tests (41.49, 20.40) supporting Stoddard's claim that receptive knowledge is more easily gained than productive. Similar results were also found in subsequent tests three days, eight days, and 28 days after the treatment.

In Waring's (1997b) study, EFL learners were presented with two sets of 15 low frequency words, one of which they learned receptively, and the other they learned productively. Subjects were given as much time as they needed, and the time it took to learn receptively and productively was recorded. The sets of words, and the order of learning were controlled to ensure that there was not a learning effect. The task was completed when the subjects were able to recall one set of words twice receptively, and the other twice productively without mistakes. After the treatment was completed there was a 10 minute distractor exercise, followed by the two translation tests. In this study, subjects completed both the receptive and the productive tests with the productive measure always administered first. Three delayed post-tests were given one day, one week, and three months after the treatment.

The results of the first tests completed after the treatment were remarkably similar to those of Griffin and Harley. The mean scores on the receptive test were slightly higher

for the subjects who completed the receptive task (12.83) than for those who completed the productive task (10.82). However, the results of the productive test showed that subjects who completed the productive task had much higher mean scores (11.20) than for those who completed the receptive task (6.07). Unfortunately, a statistical analysis was not performed between the test type and the learning types but it would appear that the difference between these mean scores would be highly significant. This suggests that if one's aim is to increase both receptive and productive knowledge, then productive learning tasks may be more effective. The extremely low scores by the receptive learners on the productive test also indicate that receptive tasks may not be well suited to developing productive knowledge. This is supported by Griffin and Harley's findings. The results also showed that the productive learning task took significantly longer (26% more time was needed) than the receptive learning task. This supports earlier findings indicating receptive knowledge may be more easily gained than productive. Furthermore, the rate of decay was much higher for productive knowledge than for receptive knowledge, contrasting with Griffin and Harley's findings. However, this may be due more to the experimental design than to greater productive decay. Since the answers to the receptive test were presented to the learners in the productive test, it is highly likely that there would have been a learning effect. On the productive test, the subjects would only have been able to draw on the knowledge that was gained in the learning task. Therefore, Griffin and Harley's findings showing a similar rate of decay for both receptive and productive learning are likely to be more accurate.

Taken as a whole, the results of Stoddard (1929), Waring (1997b) and Griffin and Harley (1996) are valuable for three reasons. First, they demonstrate that learning is bi-directional. Learners may acquire productive knowledge from receptive learning tasks and vice versa. This supports research involving incidental learning from reading which showed that learners can gain productive knowledge from a reading task (see for example, Rott, 1999; Experiment 1). Second, they also indicate that neither receptive nor productive learning tasks are ideal on their own. Receptive tasks may not be as effective at increasing productive knowledge as productive tasks while productive tasks

may be less effective at increasing receptive knowledge. This is important because it suggests to teachers and students that to have full knowledge of a word, a combination of receptive and productive learning may be needed. However, if a learner's goal is comprehension, receptive learning may be more effective, and if the goal is to improve speaking or writing, productive learning tasks may be more useful.

Third, if full knowledge of a word is desired and there is only time for one type of task, a productive task may be more useful. Receptive learning may not be well suited to developing productive knowledge while productive learning may lead to adequate gains of both receptive and productive knowledge.

One weakness of the studies is that they do not measure any aspects of knowledge besides meaning and form. Experiments 1 and 2 demonstrated that learners might gain receptive and productive knowledge of form, syntax, grammatical functions, and associations as well as meaning through learning from word pairs. Further research investigating which knowledge types are gained through receptive and productive learning is needed.

Surprisingly, there are no other studies that have investigated the effects of productive learning, and there is only one study that has examined the effects of receptive learning. Crow and Quigley (1985) investigated two different approaches to vocabulary learning to determine which promoted greater gains in receptive knowledge. The first approach involved learning words receptively in semantic fields while the second approach combined receptive and productive learning using a more 'traditional' methodology. Twice as many words (180) were studied using the receptive learning condition than using the combined approach (90). Vocabulary learning was measured with three tests after the completion of the treatment and also four weeks later. The first test was a multiple-choice test from which the subjects had to select the correct item from four choices to complete sentences. The second test was a cloze test in which the subjects were required to choose the best word from 16 choices to fill in the 15 blanks in a paragraph, and the third test was a matching test in which the subjects had to match a

word with an underlined word or phrase in a paragraph. On the post-test both groups had mean scores higher than 74 percent; however, the subjects who learnt using the combined approach had significantly higher scores than those who learnt using the receptive learning condition. On the follow-up test, no significant differences were found between the two approaches. A second follow-up test that had a different format from the earlier tests was administered 12 weeks after the treatment. It consisted of two measures. Again, the first test was multiple-choice. Subjects were required to select the best keyword for an underlined word in a sentence, and the second test involved writing words that were related to items underlined in a paragraph. The subjects who had learned using the receptive approach significantly outperformed the other subjects on the second follow up test. Crow and Quigley conclude that the receptive learning condition may be more effective than the combined learning condition because of the superior long term receptive gains and the larger number of words met while learning.

While the study is important because it is the first to specifically investigate whether receptive learning tasks are better suited to increasing receptive knowledge, there are several problems with the study. First, the tasks in the receptive learning condition may not have been strictly receptive tasks. One of the tasks involved discussing a passage in which the target words were met, and in another task, subjects had to substitute target words with related words in a passage. Although the tasks were not described in detail, both would appear to be productive. Second, productive knowledge not receptive knowledge was measured on the second test in the second follow-up testing session. It could best be described as a test of cued recall.

The above problems with the Crow and Quigley study and the high scores produced on the first post-test make it difficult to conclude that either learning condition was superior; both conditions led to large gains in receptive knowledge. Further research investigating the effects of receptive and productive learning needs to be carried out. More specifically the following questions remain to be answered:

1. Does receptive learning lead to larger gains in receptive knowledge of orthography, association, syntax, grammatical functions, and meaning and form than productive learning?
2. Does productive learning lead to larger gains in productive knowledge of orthography, association, syntax, grammatical functions, and meaning and form than receptive learning?

4.5 The present study

In the present study, three experiments were designed to examine the relative effectiveness of two productive and two receptive learning tasks: receptive and productive learning from word pairs, sentence production, and learning from glossed sentences. Receptive and productive gains in knowledge of orthography, association, syntax, meaning, and grammar were measured using the same 10 test formats in Experiment 2. It is important to note that the receptive and productive tests were not equally difficult and therefore the size of the gains for each aspect cannot be compared. It was expected that the subjects would score higher on the tests that measured receptive knowledge. The productive tests have an open format that offers little opportunity to score correctly through guessing while all but one of the receptive measures are multiple-choice tests that provide a 25 percent chance of scoring correctly without knowledge of the target words. Although the tests of meaning and form employ the same format (translation), scores may also be biased in favor of the receptive test because of the experimental design. Since the productive test was completed before the receptive test, there may be a learning effect because the answers for the receptive test were shown in the productive measure. The bias in favor of the receptive tests does not pose any problems to this study because the aim is to compare productive and receptive vocabulary learning tasks rather than vocabulary size. At present there does not appear to be a test format that would allow a comparison of the size of receptive and productive gains for any of the knowledge types except for meaning and form. Translation tests may prove to be an effective measure of vocabulary size. This would be a useful area for future research.

Another aim of this study was to find how the learning tasks affected each aspect of knowledge. It was therefore very important to isolate each type of knowledge. If a test involved more than one aspect, the gain would be unclear. It would also be uncertain how each type of knowledge influenced the scores on that test. This eliminated the possibility of using any tests that provided context to cue recall despite suggestions that context is essential on receptive tests (Crow, 1986). There were several reasons not to include context. First, providing context on tests may have resulted in a learning effect. If contexts were presented in any of the tests, the subjects may have been able to gain knowledge of a target word's part of speech, syntagmatic associations, and meaning and form which they could use in a subsequent test. Second, if contexts are provided, subjects might be able to use another type of knowledge rather than the intended aspect to answer the question. For example,

I am _____ of snakes.
a) interested b) worried c) afraid d) nervous

If this is a test of productive knowledge, it is unclear which type of knowledge is being tested because the answer involves knowledge of meaning and form, and syntax. While the determining factor in the answer is the syntagmatic association between *afraid* and *of*, a learner may also choose that response because of a semantic link between the words *snake* and *afraid*. Since the answer involves more than one knowledge type, it is unclear which type of knowledge is being tested, and which type of knowledge was used to answer the question.

A third reason not to include context on the tests was that transfer-appropriate processing theory suggests that the more a test resembles the learning condition the better learners are likely to do (Griffin, 1992). Since contexts were presented in the learning from glossed sentences task but not in the sentence production task, it would be unclear whether higher scores from the contextualized task were the result of the difference in treatments or the similarity between the testing and learning conditions.

A fourth reason is there is always a possibility that a subject may answer incorrectly not because of a lack of knowledge of the target words but because of a lack of knowledge of words presented in the context. If a subject cannot understand the context he or she may not be able to choose the correct answer for that reason.

4.6 Tests

The following section outlines the reasons the test formats were chosen. For a detailed description of the tests see Experiments 1 and 2.

All of the productive tests had similar formats in which a decontextualized cue prompted a response. The format on the tests of orthography, syntax, and association could best be described as word association. The subjects were cued with the target words (on the test of orthography the subjects were cued with a phonological prompt) and had to produce a response that demonstrated the aspect of knowledge being measured. For example, on the test of syntax, the subjects had to produce a syntagmatic associate, and on the test of association, the subjects needed to produce a paradigmatic associate. The test of grammar had a similar format, the responses were cued with target words but the responses were sentences rather than single words. It was made clear in the instructions that the determining factor for a correct response was using the target word with grammatical accuracy. The test of meaning was a productive measure of recall. On this test the cue was the L1 meaning. The formats on the productive tests were selected for the following reasons:

1. They clearly demonstrated productive knowledge.
2. The aspects of knowledge could be isolated in the tests.
3. They did not require context.
4. They were easy to understand, and quick to complete.
5. There was little difference between the formats. This allowed a comparison of the gains on the different tests.
6. The tests had proven effective in the two previous experiments.

Several other formats were also considered for the productive tests. Spew tests may be the ideal format for measuring productive knowledge (Melka, 1997); however, a weakness of spew tests is that they cannot show whether subjects fail to answer correctly because of a lack of knowledge or an inability to remember the target words. Since the determining factor might be memory, it was considered unsuitable. Sentence production was also considered for the test of syntax. However, it might be ineffective at times because it is possible to write a perfect sentence without demonstrating knowledge of syntagmatic associations. For example, for the target word *hot* the sentence 'It is hot.' demonstrates knowledge of grammar but not syntax.

More demanding measures such as those administered in studies by Schmitt and Meara (1997), and Schmitt (1998) were also considered. The test formats were well designed and effective but seem better suited to longitudinal studies. Schmitt's (1998) test of meaning and form required learners to demonstrate knowledge of the different meaning senses of target words. Productive knowledge of grammatical functions was measured by having subjects name the part of speech of the target word and then say any derivative forms of the word. Schmitt and Meara (1997) used a similar measure in which subjects had to write any possible suffixes for the target words. Schmitt and Meara (1997) and Schmitt (1998) measured knowledge of association using more demanding word association tests requiring three responses. Since they did not discriminate between paradigmatic associations and syntagmatic associations it was not clear whether they were measuring knowledge of association or syntax as defined in this study. Although the tests in the two studies effectively measured aspects of productive knowledge, they were considered inappropriate for this study because they measured fuller knowledge of the target words than the subjects could be expected to gain. Based on pilot studies and the results of Experiment 1, these tests would likely have resulted in floor effects.

The receptive tests of orthography, syntax, association, and grammar were multiple-choice. The test of meaning was a measure of receptive recall in which the target words cued responses of L1 form. An argument can be made that a receptive

translation test is both receptive and productive (Waring, 1999); however, the same argument can be made for most receptive tests. Multiple choice and matching tests were considered inappropriate for the test of meaning because their use may have resulted in a ceiling effect. Since the subjects had learned the answers in the treatments, a recognition test would have been an extremely easy task. The receptive translation test was better suited for this experiment because it was more demanding, requiring learners to recall meanings rather than recognize them. The receptive measures were selected for the following reasons:

1. The formats are commonly used to measure receptive knowledge.
2. The formats were easy to understand, and quick to complete.
3. The subjects were familiar with the multiple-choice format.
4. They do not require context.
5. The tests had proven effective in the two previous experiments.

Several other formats were considered for the receptive measures. The checklist format may have been equally effective but it did not appear to offer any particular improvement over the multiple-choice tests. The advantage of the multiple-choice tests was that the subjects were familiar with the format. Since the time taken to complete the treatment and tests, and subject fatigue were relevant factors, less familiar test formats were not selected unless they were clearly superior. Meara and Buxton (1987) demonstrated the checklist format is a reliable and valid method of assessing vocabulary knowledge. Since it also could be used to isolate a knowledge type, is relatively fast to complete, and does not require context, it may prove to be an effective method for measuring different aspects of receptive knowledge. This would be a useful area for future research.

Tests administered in the studies by Schmitt (1998) and Schmitt and Meara (1997) also may be effective in isolating and measuring receptive knowledge but would appear better suited to longitudinal studies. Schmitt (1998) measured receptive knowledge of meaning and form by giving verbal prompts to determine which meaning senses of

each target word were known. Schmitt and Meara (1997) measured receptive knowledge of grammar by having subjects circle allowable suffixes for each target word. While both tests effectively isolated the desired knowledge type they were considered too demanding for this study.

4.7 Implementation of the research

4.7.1 Research Questions

The present study will address the following questions:

1. How do receptive and productive learning tasks affect vocabulary knowledge?

Research involving word pairs suggests that receptive and productive learning are bi-directional (Griffin and Harley, 1996; Stoddard, 1929; Waring, 1997b). However, the only knowledge type that was tested in those studies was meaning and form. Experiments 1 and 2 indicate that knowledge of association, syntax, grammatical functions, and orthography may also be gained through receptive learning. It is likely that those aspects of knowledge may also be gained from productive learning tasks.

2. Which is more effective, receptive or productive learning?

Research on word pairs suggests that receptive tasks are more effective in increasing receptive knowledge while productive learning is more effective in promoting productive knowledge (Griffin and Harley, 1996; Stoddard, 1929; Waring, 1997b). Griffin and Harley state that if only one task is used, productive tasks may be more effective in developing fuller vocabulary knowledge. Measuring different types of receptive and productive tasks with

10 tests of vocabulary knowledge may provide more evidence to support this claim.

4.7.2 Pilot studies

Pilot studies were conducted with Japanese learners studying English as a foreign language to determine the following:

1. Subjects needed approximately six minutes to complete the treatments in Experiment 3, 12 minutes in Experiment 4, and a maximum of 20 minutes in Experiment 5.
2. The subjects were likely to know all of the running words in the contexts.
3. The writing task was likely to be more effective than the reading task.

4.8 Experiment 3

4.8.1 Subjects

The subjects in this experiment were 62 Japanese native speakers from two first-year EFL classes at Kyushu University in Fukuoka, Japan. All of the subjects had studied English for a minimum of six years, and scored 80 percent or higher on Version 1 of the Vocabulary Levels Test at the 2000 word level. Their mean score was 27.5 indicating that they were well in control of that level, and that they had receptive knowledge of almost all of the 2000 most frequent words (Schmitt, Schmitt, and Clapham, 2001). Their mean score on version C of the Productive Vocabulary Levels Test at the 2nd 1000 word level was 14.7 suggesting that they had productive knowledge of approximately 1700 of the 2000 most frequent words. The subjects were randomly assigned to the experimental groups.

4.8.2 Design

Two experimental groups were used in this study to examine the effects of receptive and productive learning from word pairs. One experimental group studied 10 L2 target words receptively, and the other studied the same target words productively. In the receptive treatment, the target words were presented in a column on the left side of a paper, and their translations were presented on their right (see Appendix G). The subjects were instructed to cover the translations and then look at the target words and try to recall their translations. If they could not recall a meaning, the subjects were told to uncover the translation and check the meaning. In the productive treatment, the translations were on the left and the target words were on the right (see Appendix H). The subjects completed the same task except they covered the target words, and then tried to recall them. Each group was given six minutes to complete their task.

After the completion of the treatments, ten tests measuring vocabulary knowledge were administered. The tests used the same formats as those used in Experiment 2, and the subjects completed the same procedure used in Experiments 1 and 2 when completing the tests. The results of the groups were then compared to determine how receptive and productive learning from word pairs contributed to vocabulary learning.

4.8.3 Procedure

The experiment was conducted in three stages over two weeks. The first stage of the experiment was mandatory for all students in the classes while the second and third stages were optional. In the first stage, students completed the receptive and productive Vocabulary Levels Tests at the 2000 word level. In the second stage of the experiment, two subjects who did not take part in stage three were involved in case studies outside of regular class time. The subjects were randomly assigned to the experimental groups. While they completed the treatment and dependent measures they were asked to think out loud to provide insight into subjects' thought processes

during the experiment. Their results were not included in stage three. During the third stage of the experiment, the students completed the treatments and tests.

The experimental procedure was implemented over the following three stages:

- Stage 1 The Vocabulary Levels Tests were administered to determine the vocabulary size of potential subjects. Subjects were randomly assigned to treatment groups.
- Stage 2 Case studies were conducted.
- Stage 3 Subjects participated in treatments, and the dependent measures were administered.

At the beginning of stage 3, the instructor explained and demonstrated the two tasks. The subjects were told that they would have six minutes to learn the 10 target words, and would be tested after the treatment. Any questions were addressed to make certain that the instructions were completely understood. The subjects were not given any information about how they would be tested. Each task was presented on one page with instructions written in Japanese at the top of the page. During the treatment, the instructor monitored the subjects to ensure that they were using the desired learning tasks.

After the treatments, the subjects were given 10-page test booklets. There was one test on each page, and they were instructed to complete them in order. Once a test was finished they were not allowed to go back to it and revise any answers. They were carefully supervised to ensure that this did not occur. This was to eliminate the possibility of a learning effect from subsequent tests. Subjects had as much time as they needed to finish a test.

4.8.4 Materials

4.8.5 Target words

Ten target words were used in this experiment. The target words were the same 10 words with high frequency synonyms used in Experiment 2 (Set A). The target words were replaced with the same ten nonsense words created and used in Experiment 1 to ensure that the subjects had no prior knowledge of the target words.

4.8.6 Time

Subjects were given six minutes to complete the treatments. It was determined in pilot studies that six minutes was sufficient for the subjects to finish their tasks.

4.8.7 Dependent Measures

The same ten tests that were administered in Experiment 2 were used to measure the effects of receptive and productive learning tasks on vocabulary knowledge in Experiments 3, 4, and 5. Each test enabled learners to demonstrate a specific aspect of word knowledge productively or receptively. The tests measured receptive and productive knowledge of orthography, meaning and form, grammatical functions, association, and syntax.

4.8.8 Case studies

Case studies were done with two learners. Each subject completed one of the treatments. In the case studies, the subjects were asked to think aloud during both the treatment and tests. After the completion of the tests, questions were asked to gain further insight into the learning process.

The following observations emerged from the case studies:

1. The subjects believed that they had successfully learnt meaning and form in the treatments. However, after completing the two spelling tests some forgetting did occur.
2. The subject that completed the productive task used mnemonics to learn the spelling of the target words.
3. The subjects initially attempted to use L1 knowledge in the productive measures of association, and syntax. The subject that completed the productive task also considered L2 knowledge of synonyms on the test of syntax if she doubted the transferability of an L1 syntagmatic associate.
4. The subject that completed the productive task considered the form of the target words first to determine their part of speech. If she was still unsure, she based her answers on the tests of grammar on L1 grammatical knowledge. The receptive subject based all of her answers on L1 grammatical knowledge.
5. Both subjects looked for information in the distractors on the receptive tests to confirm or discount answers. This strategy worked well for both of them.

4.8.9 Results

The descriptive statistics (means, standard deviations, and number of subjects) of vocabulary knowledge scores for the 10 dependent measures are reported in Table 23. To determine whether there were any overall differences among the treatment groups, a multivariate analysis of variance (MANOVA) was performed using the scores on the 10 dependent measures (productive orthography, receptive orthography, productive meaning and form, receptive meaning and form, productive association, receptive association, productive syntax, receptive syntax, productive grammar, and receptive grammar). The independent variable was the type of learning task (receptive and

productive word pairs). The MANOVA revealed an overall significant difference (Pillai's Trace, Hotelling's Trace, Wilks' Lambda, and Roy's Greatest Root) between the two tasks ($F(10,51) = 3.73, p < .001$).

Table 23 Means and standard deviations on dependent measures for the receptive and productive word pairs tasks

| Learning Condition | <i>n</i> | PO | RO | PM | RM | PA | RA | PS | RS | PG | RG |
|-----------------------|----------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| Productive word pairs | 29 | 8.79 (1.50) | 9.24 (0.95) | 8.79 (1.93) | 7.83 (3.13) | 8.10 (2.51) | 8.48 (1.92) | 6.21 (2.47) | 8.03 (1.95) | 8.76 (1.79) | 8.79 (1.80) |
| Receptive word pairs | 33 | 7.76 (1.77) | 8.61 (1.27) | 7.67 (2.19) | 9.09 (1.61) | 7.15 (2.29) | 8.73 (1.53) | 4.82 (2.30) | 8.27 (1.81) | 7.30 (2.38) | 9.00 (1.12) |

Note. Maximum score = 10, standard deviations are in parentheses, Key. PO=productive knowledge of orthography, RO=receptive knowledge of orthography, PM=productive knowledge of meaning and form, RM=receptive knowledge of meaning and form, PA=productive knowledge of association, RA=receptive knowledge of association, PS=productive knowledge of syntax, RS=receptive knowledge of syntax, PG=productive knowledge of grammar, RG=receptive knowledge of grammar.

Table 24 Summary of multivariate analysis of variance on vocabulary tests in Experiment 3

| | productive tests | receptive tests |
|-------------|---------------------------------|------------------------------|
| orthography | $F(1,60) = 6.10, p = .016^*$ | $F(1,60) = 4.84, p = .032^*$ |
| meaning | $F(1,60) = 4.55, p = .037^*$ | $F(1,60) = 4.14, p = .046^*$ |
| association | $F(1,60) = 2.43, p = .124$ | $F(1,60) = .311, p = .579$ |
| syntax | $F(1,60) = 5.26, p = .025^*$ | $F(1,60) = .249, p = .620$ |
| grammar | $F(1,60) = 7.26, p = .009^{**}$ | $F(1,60) = .303, p = .584$ |

Note. $^*p < .05$, $^{**}p < .01$, $^{***}p < .000$

Table 23 shows that the subjects that learned the target words in the productive task outperformed those that completed the receptive task on all of the productive measures and also on the receptive orthography test. Significant differences were found on the productive tests of orthography ($F(1,60) = 6.10, p < .05$), meaning and form ($F(1,60) = 4.55, p < .05$), syntax ($F(1,60) = 5.26, p < .05$), grammar ($F(1,60) = 7.26, p < .01$), and the receptive test of orthography ($F(1,60) = 4.84, p < .05$). In contrast, the receptive word pairs subjects demonstrated larger gains on all of the receptive measures except for the receptive test of orthography. However, only the gains on the receptive test of meaning

were significantly higher ($F(1,60) = 4.14, p < .05$). A summary of the statistical analysis is shown in Table 24.

The results support previous findings that suggest that the type of learning task (receptive or productive) will influence the type (receptive and productive) and amount of knowledge gained (Stoddard 1929; Griffin and Harley 1996; Waring 1997b). In the earlier studies only knowledge of meaning was measured. The results of this experiment suggest that knowledge of association, syntax, and grammar may also be affected in the same way. Receptive learning is likely to lead to higher gains in receptive knowledge than productive knowledge while for productive learning the opposite is true. The reason significant differences were not realized on the receptive knowledge of association, syntax, and grammar measures may be due to ceiling effects on those tests. The results suggest that all of the receptive tests may have been too sensitive to gains in knowledge. Further research is needed to investigate whether receptive learning leads to significantly larger gains in receptive knowledge for those knowledge types.

Table 25 Means and standard deviations of combined receptive and productive scores for each aspect of knowledge in Experiment 3

| Learning Condition | <i>n</i> | Orthography | Meaning | Association | Syntax | Grammar |
|-----------------------|----------|--------------|--------------|--------------|--------------|--------------|
| Productive word pairs | 29 | 18.03 (2.06) | 16.62 (4.11) | 16.59 (3.59) | 14.24 (3.86) | 17.55 (3.34) |
| Receptive word pairs | 33 | 16.36 (2.73) | 16.76 (3.25) | 15.88 (3.56) | 13.09 (3.60) | 16.30 (3.20) |

Note. Maximum score = 20, standard deviations are in parentheses

Table 25 shows the combined scores on the receptive and productive tests for each aspect of knowledge. A second MANOVA was performed to determine if there were any significant differences between any of the combined scores for each of the five aspects. The MANOVA revealed an overall significant difference (Pillai's Trace, Hotelling's Trace, Wilks' Lambda, and Roy's Greatest Root) between the two tasks ($F(5,56) = 3.42, p < .01$). In the analyses of the differences between the two tasks for each aspect, the only significant difference that was found was for orthography

($F(1,60) = 7.25, p < .01$). The productive word pairs task led to significantly higher gains in knowledge ($p < .01$) than the receptive word pairs task. This may be due to the focus in productive learning on the form of the target word. In the receptive task, learners are more focused on L1 meaning. The overall scores for meaning were almost identical for learners from both groups, which is similar to Stoddard's (1929) and Griffin and Harley's (1996) findings. The productive learning task led to combined scores for association, syntax, and grammar that were all higher than the corresponding receptive scores. A summary of the results of the MANOVA are shown in Table 26.

Table 26 Summary of multivariate analysis of variance on combined receptive and productive scores for each aspect of knowledge

| | |
|-------------|---------------------------------|
| orthography | $F(1,60) = 7.25, p = .009^{**}$ |
| meaning | $F(1,60) = .021, p = .884$ |
| association | $F(1,60) = .605, p = .440$ |
| syntax | $F(1,60) = 1.47, p = .230$ |
| grammar | $F(1,60) = 2.24, p = .140$ |

Note. $^{**}p < .01$

Taken as a whole, the results suggest that both types of learning are very effective in developing vocabulary knowledge. Both groups demonstrated large gains in all of the knowledge types. The results supported previous findings that examined how receptive and productive learning affect meaning. Receptive learning might be more effective in developing receptive knowledge of meaning, grammar, association and syntax while productive knowledge may lead to larger gains in productive knowledge for those aspects as well as both receptive and productive knowledge of orthography.

If only one type of learning is used the productive word pairs task may be superior to the receptive word pairs task. The combined scores show that the productive learning task led to higher gains in knowledge of all aspects of knowledge except meaning. Productive learning led to significantly larger gains on all of the productive measures except association as well as the receptive test of orthography. Although subjects that completed the receptive task significantly outperformed the productive subjects on the

receptive test, differences were not significant on the receptive measures of association, syntax, and grammar.

4.9 Experiment 4

Experiment 4 was undertaken to determine how other receptive and productive tasks contributed to vocabulary knowledge. Specifically, the experiment sought to examine the relative effectiveness of learning from three glossed sentences, and learning from sentence production.

4.9.1 Subjects

The subjects in this experiment were 66 Japanese native speakers from one first-year EFL class at Kyushu University in Fukuoka, Japan. Their mean score on Version 1 of the Vocabulary Levels Test at the 2nd 1000 word level was 27.6 indicating that they were well in control of that level, and that they had receptive knowledge of almost all of the 2000 most frequent words (Schmitt, Schmitt, and Clapham, 2001). Their mean score on version C of the Productive Vocabulary Levels Test at the 2nd 1000 word level was 15.1 suggesting that they may have mastered that level and had productive knowledge of approximately 1800 of the 2000 most frequent words (Laufer and Nation, 1999). The subjects were randomly assigned to the experimental groups.

4.9.2 Design

Two experimental groups took part in this study to examine the effects of receptive vocabulary learning from reading and productive vocabulary learning from writing. Subjects in the receptive treatment met each target word in word pairs and in three glossed sentences (see Appendix I). The same 10 target words and nonsense words from Experiment 3 were used in this experiment. The target words were presented on the left of their L1 translations followed by the sentences. In each sentence, the target

words were underlined and written in bold. In the following example the meaning of the target word ‘dangy’ is *boulder*.

dangy 巨石 The **dangy** was as large as a small house.
On the way up the mountain we passed a **dangy**.
He stood on the biggest **dangy** to get a better view.

In the productive treatment, the target words were presented in the same word pairs followed by space to write each target word in a sentence (see Appendix J). For example,

dangy 巨石 _____

Instructions on how to complete each treatment were written in the subjects’ L1 at the top of each page. The subjects were given 12 minutes to complete the treatments. It was determined in pilot studies that 12 minutes would allow the slowest students to finish the productive task. The same 10 tests from the previous experiment were administered after the treatments (for an example of the test formats see Appendix F). The subjects were told that they would be tested after the treatments but they were not told the nature of the tests.

4.9.3 Sentence contexts

The contexts were from the British National Corpus. Since the amount of information a context provides about a target word was found to influence learning gains (Experiment 1), an attempt was made to select equally informative contexts. The level of information they provided about the target words could be described as follows: It is unlikely that the exact meaning of the target word can be inferred. However, information in the context may lead to partial knowledge of the target word’s meaning. The criteria for selecting the contexts were described in the methodology section of Experiment 2.

Extensive pilot testing was used to ascertain that subjects were familiar with all of the running words in the contexts. Contexts that contained words that were unknown to any subject in the pilot tests were either simplified or discarded. Controlling the running words in the contexts ensured that the learners would quickly understand the sentences. It also eliminated the chance of a learning reduction because of other unknown words in the contexts.

4.9.4 Case studies

Case studies were done with two learners, each completing one of the treatments. Several observations emerged from the two case studies.

1. The subject that completed the reading task attempted to memorize both the target words and their L1 meanings, and also the sentences during the treatment. The subject that completed the writing task reread the sentences that she wrote after finishing the task and considered revising them.
2. In the reading task, the subject paid little attention to the spelling of the target words and made several spelling mistakes on the productive measure of orthography by combining the spelling of two target words. The subject in the writing task was more careful to learn the spelling of the target words.
3. The subject that completed the reading task had very little problem with all of the tests. The subject that completed the writing task confused two answers on the productive tests of meaning, syntax, and association. However, he was able to score well on all of the productive tests, and score correctly on all of the receptive tests.

4.9.5 Results

The descriptive statistics (means, standard deviations, and number of subjects) of vocabulary knowledge scores for the 10 dependent measures are reported in Table 27.

A multivariate analysis of variance (MANOVA) was performed using the scores on the 10 dependent measures. The independent variable was the type of learning task (writing each target word in a sentence and reading each target word in three sentences). The MANOVA revealed an overall significant difference (Pillai's Trace, Hotelling's Trace, Wilks' Lambda, and Roy's Greatest Root) between the two tasks ($F(10,55) = 5.07, p < .000$).

Table 27 Means and standard deviations of learning conditions on dependent measures in Experiment 4

| Learning Condition | <i>n</i> | PO | RO | PM | RM | PA | RA | PS | RS | PG | RG |
|-------------------------|----------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| Writing a sentence | 31 | 8.13 (2.05) | 8.74 (1.34) | 7.61 (1.84) | 8.84 (1.59) | 7.32 (1.89) | 8.84 (1.19) | 5.94 (2.54) | 8.32 (1.30) | 8.39 (1.75) | 9.03 (1.30) |
| Reading three sentences | 35 | 9.17 (1.46) | 9.49 (0.89) | 9.43 (1.12) | 9.09 (1.84) | 8.71 (1.47) | 9.06 (1.26) | 7.29 (2.88) | 8.74 (1.40) | 9.06 (1.73) | 9.54 (0.98) |

Note. Maximum score = 10, standard deviations are in parentheses Key. PO=productive knowledge of orthography, RO=receptive knowledge of orthography, PM=productive knowledge of meaning and form, RM=receptive knowledge of meaning and form, PA=productive knowledge of association, RA=receptive knowledge of association, PS=productive knowledge of syntax, RS=receptive knowledge of syntax, PG=productive knowledge of grammar, RG=receptive knowledge of grammar.

Table 27 shows that the subjects that completed the receptive task outperformed the productive group on all of the dependent measures. Significant differences were found on the productive tests of orthography ($F(1,64) = 5.76, p < .05$), receptive orthography ($F(1,64) = 7.21, p < .01$), productive meaning ($F(1,64) = 24.10, p < .000$), productive association ($F(1,64) = 11.33, p < .01$), and productive syntax ($F(1,64) = 4.04, p < .05$). A summary of the results of the statistical analysis is shown in Table 28.

Table 28 Summary of multivariate analysis of variance on vocabulary tests

| | productive tests | receptive tests |
|-------------|-----------------------------------|---------------------------------|
| orthography | $F(1,64) = 5.76, p = .019^*$ | $F(1,64) = 7.21, p = .009^{**}$ |
| meaning | $F(1,64) = 24.10, p = .000^{***}$ | $F(1,64) = .336, p = .564$ |
| association | $F(1,64) = 11.33, p = .001^{**}$ | $F(1,64) = .522, p = .472$ |
| syntax | $F(1,64) = 4.04, p = .049^*$ | $F(1,64) = 1.58, p = .213$ |
| grammar | $F(1,64) = 2.44, p = .123$ | $F(1,64) = 3.28, p = .075$ |

Note. $^*p < .05$, $^{**}p < .01$, $^{***}p < .000$

Table 29 shows the combined receptive and productive scores for each aspect. A second MANOVA revealed an overall significant difference (Pillai's Trace, Hotelling's Trace, Wilks' Lambda, and Roy's Greatest Root) between the two tasks ($F(5,60) = 2.51, p < .05$) for the combined scores. Significant differences were found for orthography ($F(1,64) = 8.28, p < .01$), meaning ($F(1,64) = 8.44, p < .01$), association ($F(1,64) = 6.46, p < .05$), and syntax ($F(1,64) = 4.29, p < .05$). However, no significant difference was found for grammar. A summary of the statistical findings is shown in Table 30.

Table 29 Means and standard deviations of combined receptive and productive scores for each aspect of word knowledge in Experiment 4

| Learning Condition | <i>n</i> | Orthography | Meaning | Association | Syntax | Grammar |
|--------------------|----------|--------------|--------------|--------------|--------------|--------------|
| Writing | 31 | 16.87 (2.81) | 16.45 (3.09) | 16.16 (2.67) | 14.26 (3.25) | 17.42 (2.68) |
| Reading | 35 | 18.66 (2.22) | 18.51 (2.68) | 17.77 (2.47) | 16.03 (3.65) | 18.60 (2.51) |

Note. Maximum score = 20, standard deviations are in parentheses

Table 30 Summary of multivariate analysis of variance on combined receptive and productive scores for each aspect of knowledge

| | |
|-------------|---------------------------------|
| orthography | $F(1,64) = 8.28, p = .005^{**}$ |
| meaning | $F(1,64) = 8.44, p = .005^{**}$ |
| association | $F(1,64) = 6.46, p = .013^*$ |
| syntax | $F(1,64) = 4.29, p = .042^*$ |
| grammar | $F(1,64) = 3.41, p = .069$ |

Note. $^*p < .05$, $^{**}p < .01$, $^{***}p < .000$

Overall, the results indicate that both tasks were very effective. Large gains in knowledge of all aspects were made by both groups. However, the results show that the receptive task is superior to the productive task on both the productive and receptive measures. This contrasts with previous research on word pairs, and Experiment 3. One factor that may have contributed to this result was time. The amount of time given (12 minutes) was based on how much time the slowest subjects needed to complete the productive writing task. In pretests, it was found that most students needed only eight minutes to complete the reading task. Thus, subjects in the receptive group may have used alternative strategies such as learning from word pairs receptively and productively or creating mnemonics to aid them in their learning. While some subjects in the productive group may have also used other strategies, the amount of time they had on average would have been less than was available to the receptive group. To clarify what is gained from each task a fifth experiment was performed.

4.10 Experiment 5

In the previous experiment, subjects were given an equal amount of time to complete the treatments. That might have allowed subjects in the receptive treatment and to a lesser degree in the productive treatment the time to use other strategies to learn the target words. The next experiment was designed to clarify the previous findings, and determine how each task affects vocabulary knowledge. There were two primary differences between Experiments 4 and 5. First, in Experiment 4 the subjects knew that they would be tested at the conclusion of the treatments. Since, this may have motivated them to use other learning strategies during the treatments, the subjects in Experiment 5 were not told that they would be tested. Second, in Experiment 5, as soon as a subject completed both tasks they were finished. Time was considered to be a function of the tasks, and was not controlled. Therefore, it was likely that time on tasks, and on target words would covary. Hulstijn and Laufer (2001) took a similar approach arguing that time should not be considered a separate variable but an inherent property of a task.

4.10.1 Subjects

The subjects in this experiment were 49 Japanese native speakers from one first-year EFL class at Kyushu University in Fukuoka, Japan. Their mean score on Version 1 of the Vocabulary Levels Test at the 2nd 1000 level was 27.1 indicating that they were well in control of that level, and that they had receptive knowledge of almost all of the 2000 most frequent words (Schmitt, Schmitt, and Clapham, 2001). Their mean score on version C of the Productive Vocabulary Levels Test at the 2nd 1000 word level was 14.8 suggesting that they had productive knowledge of approximately 1750 of the 2000 most frequent words. The subjects were randomly assigned to the experimental groups.

4.10.2 Design

Experiment 5 involved the same treatments as Experiment 4 but differed from the previous design in three respects. First, in Experiment 5, one experimental group completed both treatments. This reduced the number of subjects required, and also decreased the possibility that individual differences might influence the results. Second, in Experiment 4, the subjects were given an equal amount of time (12 minutes) to complete their tasks; however, in Experiment 5 the subjects were only given enough time to complete each task. The subjects were instructed to begin the second treatment as soon as they had completed the first one. When the subjects finished the second treatment, they turned over their treatment booklets, which were then collected. That made it less likely that the subjects could use alternative learning strategies when completing the tasks. Third, the subjects did not know that they would be tested in Experiment 5. That also reduced the chance that the subjects would use alternative learning strategies when completing the tasks.

4.10.3 Target words

20 target words were used in this experiment. The target words were the same 20 words used in Experiment 2 (Set A and B). The target words were replaced with the

same 20 nonsense words used in Experiment 2 to ensure that the subjects had no prior knowledge of the target words.

Four versions of the treatment were created to ensure that interlexical and intralexical factors, and order of the tasks did not influence the results. Half of the subjects completed the productive task first, and the other half completed the receptive task first. Half of the subjects did the receptive and productive tasks with Set A of the target words, and the other half did the tasks with Set B. One of the four treatments used in Experiment 5 is shown in Appendix K. After the completion of the treatments, ten tests measuring vocabulary knowledge were administered. The same 10 test formats from Experiments 2, 3, and 4 were used in Experiment 5 (for an example of the test formats see Appendix F).

4.10.4 Sentence contexts

The contexts were from the British National Corpus. The same criteria as the previous experiments were used to select the contexts.

4.10.5 Case studies

Case studies were done with two learners. One completed the reading task first, and the other completed the writing task first. Several observations emerged from the case studies.

1. It took the subjects an average of nine minutes and 12 seconds to complete the writing task, and seven minutes and five seconds to complete the reading tasks.
2. Both subjects paid more attention to spelling in the writing task.
3. The subjects had great difficulty recalling meaning and form for target words from both tasks. They both tried to recall a complete context in order to produce the meaning.

4. Both subjects were unable to respond to most answers on the productive tests of grammatical functions, syntax, and association because they were unable to recall the meaning of the target words.
5. Both subjects tried to use information in the distractors to answer the receptive tests.
6. Overall, both subjects believed the writing task was more effective in learning the words.

4.10.6 Results

The descriptive statistics (means, standard deviations, and number of subjects) of vocabulary knowledge scores for the 10 dependent measures are reported in Table 31. A Repeated Measures MANOVA was performed using the scores on the 10 dependent measures to analyze the within subjects factor. The independent variable was the type of learning task (writing each target word in a sentence and reading each target word in three sentences). The MANOVA revealed an overall significant difference (Pillai's Trace, Hotelling's Trace, Wilks' Lambda, and Roy's Greatest Root) between the two tasks ($F(10,39) = 2.79, p < .05$).

Table 31 Means and standard deviations on dependent measures for the reading and writing tasks in Experiment 5

| Learning Condition | PO | RO | PM | RM | PA | RA | PS | RS | PG | RG |
|--------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| writing | 3.59 (2.14) | 6.45 (2.35) | 3.73 (2.63) | 4.45 (2.65) | 3.35 (2.31) | 6.12 (2.28) | 2.45 (2.45) | 5.84 (2.36) | 4.08 (2.53) | 6.69 (2.21) |
| reading | 2.63 (2.13) | 5.29 (2.08) | 2.08 (2.36) | 2.69 (2.26) | 1.98 (2.04) | 4.88 (1.92) | 1.53 (1.72) | 5.02 (2.42) | 2.78 (2.41) | 5.86 (2.48) |

Note. Maximum score = 10, standard deviations are in parentheses, n=49 Key. PO=productive knowledge of orthography, RO=receptive knowledge of orthography, PM=productive knowledge of meaning and form, RM=receptive knowledge of meaning and form, PA=productive knowledge of association, RA=receptive knowledge of association, PS=productive knowledge of syntax, RS=receptive knowledge of syntax, PG=productive knowledge of grammar, RG=receptive knowledge of grammar.

Table 31 shows that the productive writing group significantly outperformed the receptive reading group on all measures. The repeated measures MANOVA realized significant differences below $p < .000$ on the productive tests of meaning ($F(1,48) = 19.53$), and association ($F(1,48) = 14.99$), and the receptive test of meaning ($F(1,48) = 24.23$). A significant difference below $p < .01$ was found on the productive tests of orthography ($F(1,48) = 7.62$), syntax ($F(1,48) = 10.92$), and grammar ($F(1,48) = 11.86$), and the receptive measures of orthography ($F(1,48) = 9.57$), and association ($F(1,48) = 12.52$). A significant difference below $p < .05$ was realized on the receptive tests of grammar ($F(1,48) = 6.90$), and syntax ($F(1,48) = 4.18$). A summary of the findings is shown in Table 32.

Table 32 Summary of multiple analysis of variance on vocabulary tests

| | Productive tests | Receptive tests |
|-------------|-----------------------------------|-----------------------------------|
| orthography | $F(1,48) = 7.62, p = .008^{**}$ | $F(1,48) = 9.57, p = .003^{**}$ |
| meaning | $F(1,48) = 19.53, p = .000^{***}$ | $F(1,48) = 24.23, p = .000^{***}$ |
| association | $F(1,48) = 14.99, p = .000^{***}$ | $F(1,48) = 12.52, p = .001^{**}$ |
| syntax | $F(1,48) = 10.92, p = .002^{**}$ | $F(1,48) = 4.18, p = .047^{*}$ |
| grammar | $F(1,48) = 11.86, p = .001^{**}$ | $F(1,48) = 6.90, p = .012^{*}$ |

Note. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .000$

The combined receptive and productive scores for each aspect are shown in Table 33. A Repeated Measures MANOVA revealed an overall significant difference (Pillai's Trace, Hotelling's Trace, Wilks' Lambda, and Roy's Greatest Root) between the two tasks ($F(5,44) = 5.44, p < .01$) for the combined scores. The scores indicate that the writing task is the more effective task but it should also be noted that both conditions resulted in considerable learning in a short period of time. Interestingly, the tests of meaning resulted in the lowest combined scores. This suggests that studies that measure only meaning may be unable to find a significant result when it has in fact, occurred. It should also be pointed out that if only knowledge of meaning was the criterion for learning, then the productive task was nearly twice as effective as the receptive task. A summary of the statistical analysis for the combined scores is shown in Table 34.

Table 33 Means and standard deviations of combined receptive and productive scores for each aspect of word knowledge in Experiment 5

| Learning Condition | Orthography | Meaning | Association | Syntax | Grammar |
|--------------------|--------------|-------------|-------------|-------------|--------------|
| Writing | 10.04 (3.90) | 8.18 (5.04) | 9.47 (4.29) | 8.29 (3.86) | 10.78 (4.39) |
| Reading | 7.92 (3.65) | 4.78 (4.45) | 6.86 (3.65) | 6.55 (3.68) | 8.63 (4.37) |

Note. Maximum score = 20, standard deviations are in parentheses, n=49

Table 34 Summary of multiple analysis of variance on combined receptive and productive scores for each aspect of knowledge

| | |
|-------------|-----------------------------------|
| orthography | $F(1,48) = 12.48, p = .001^{**}$ |
| meaning | $F(1,48) = 24.27, p = .000^{***}$ |
| association | $F(1,48) = 17.34, p = .000^{***}$ |
| syntax | $F(1,48) = 8.81, p = .005^{**}$ |
| grammar | $F(1,48) = 11.64, p = .001^{**}$ |

Note. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .000$

4.11 Discussion

As mentioned earlier, one purpose of this study was to compare the relative effectiveness of receptive and productive vocabulary learning. The results of Experiments 3 and 5 indicate that productive tasks are more effective while the results of Experiment 4 suggest receptive tasks might be superior.

Experiment 3 demonstrated that productive learning from word pairs enabled learners to gain significantly greater knowledge of both receptive and productive knowledge of orthography, and productive knowledge of meaning, syntax, and grammar. Mean scores were also higher on the productive knowledge of association test; however, the difference was not significant. In contrast, the receptive learners outperformed the productive learners on all of the receptive measures although the only significant difference was on the receptive test of meaning. The absence of significant differences on those receptive measures might be due to a ceiling effect on those tests. A

replication of this experiment with less sensitive receptive measures or a larger number of target words may show the receptive task to be significantly better on the receptive tests of syntax, grammar, and association. That would be a useful follow up to this study. Overall, the findings indicate that both tasks were very effective, time efficient methods of gaining vocabulary knowledge. Large gains in receptive and productive knowledge were made for all aspects

The results of Experiment 3 support previous research that has shown that receptive learning from word pairs leads to larger gains in receptive knowledge of meaning while productive learning from word pairs leads to larger gains in productive knowledge of meaning (Stoddard 1929; Griffin and Harley 1996; Waring 1997b). The superiority of the productive task on all of the productive measures and the receptive test of orthography suggests that if only one task were to be used, it may be the more effective method; however, this may be slightly misleading. Productive learning is certainly better suited to developing productive knowledge, and knowledge of orthography but there is a lot of evidence that suggests that receptive learning from word pairs is better suited to gaining receptive knowledge. It has led to significantly greater gains in receptive knowledge of meaning in this experiment and in prior studies, and certainly meaning is the most important aspect to be gained. The size of the gains made by the receptive group on the receptive measures of syntax, association, and grammar, and the fact that the receptive group outperformed the productive group on all of those measures suggests that it might be the better method if the aim is to acquire receptive knowledge.

Experiment 4 examined vocabulary learning from different receptive (reading three glossed sentences) and productive (sentence production) tasks. The results indicate that both tasks were very effective methods of gaining receptive and productive knowledge of all aspects. However, the results contrast with earlier findings. Subjects who completed the reading task outperformed the writing group on all 10 dependent measures. Significant differences were found on the productive measures of orthography, syntax, meaning, and association, and also on the receptive test of

orthography. The superiority of the receptive task on the productive measures and not on the receptive measures is puzzling. Previous research indicated that if the reading task were more effective, it would be on the receptive tests. There are several possible explanations for this result. The first is that if receptive knowledge is easier to gain than productive knowledge as Waring (1997b), and Stoddard (1929) suggest, then both tasks may have allowed the learners to gain enough knowledge of the target words to score well on the receptive measures. The multiple choice format on the receptive tests of syntax, grammar, and association may have also contributed to the very high scores since they allowed the subjects at least a 25 percent chance to score correctly by guessing while the productive tests did not. Second, ceiling effects on the receptive tests may have reduced the possibility of finding significant differences. A third possible explanation is that the experimental design confounded the results. In the treatment, both groups were given the same amount of time (12 minutes) to complete their tasks but while some subjects that did the writing task struggled to finish on time, subjects from the receptive group had ample time to finish the reading task. While this was a function of the tasks – writing usually takes longer than reading – it may also be a function of receptive and productive learning in general. Waring (1997b) found that students took longer to learn from word pairs productively than receptively. Any extra time may have allowed the receptive group to process the sentences more deeply or use alternative strategies to learn the target words. Since the receptive task involved greater input (three sentences) than the productive task (one sentence), the extra time may have allowed the subjects in the receptive task to process greater input more deeply. This may have helped them to score higher on the productive measures. Moreover, if the receptive subjects used other learning strategies, it cannot be clear how the reading task affected the outcome. Still the results do indicate that if learners spend an equal amount of time on tasks, a receptive task or combination of tasks may be superior to a productive task. Moreover, this experiment shows that receptive learning tasks may not only be better at developing receptive knowledge but that they might also lead to significantly greater gains in productive knowledge.

The purpose of Experiment 5 was to clarify the results of Experiment 4, and get a better understanding of how receptive and productive learning tasks contribute to vocabulary knowledge. It involved the same reading and writing tasks as Experiment 4 but in this case the time taken to complete each task was not controlled. The results of Experiment 5 showed that subjects gained significantly more knowledge on all 10 dependent measures from the sentence production task. While this result supports previous research showing productive learning to be more effective in promoting productive knowledge, it also contrasts with previous research as well as the results of Experiment 3 that indicate that receptive learning is more conducive to gaining receptive knowledge. It can be argued that the productive task was more effective on the receptive measures because the learners were likely to have spent more time on the productive task than on the receptive task. Experiment 4 demonstrated that when the time of both tasks is equal the sentence production task might be inferior. While the results of Experiment 4 are interesting and useful, one must question whether that receptive treatment reflects authentic vocabulary learning. The tasks in Experiment 5 are certainly ecologically valid – both tasks are often used inside and outside the classroom. Teachers usually give students one or more examples of a new vocabulary item in context to try to teach a word, and many student dictionaries provide learners with three examples of an item in context. Teachers may also ask students to write new vocabulary in sentences, and students regularly write new words in lists with their meanings, and in context in order to learn them. If we consider the time spent on the tasks to be a function of the tasks, and we consider the results to be ecologically valid, then writing a sentence may be a much more effective method of gaining vocabulary knowledge than reading three sentences. Moreover, this experiment indicates that productive learning is not just superior in developing productive knowledge but can also lead to larger gains in receptive knowledge.

Combined, the results of Experiments 3 and 5 provide a powerful argument that teachers should encourage productive learning of vocabulary if the aim is to develop both receptive and productive knowledge. Common teaching practices of defining or translating an unknown word, and giving several examples of it in context may only be

conducive to developing students' receptive knowledge. Teachers and students need to be aware that using either productive tasks or a combination of receptive and productive tasks is likely to be much more effective in increasing overall vocabulary knowledge. Tasks such as productive learning from word pairs, and sentence production are two methods that could be incorporated into a vocabulary-learning program with little effort. In turn, integrating more productive learning tasks into the classroom could also narrow the gap between the size of learners' receptive and productive vocabulary. This study also demonstrates that receptive learning can be a very effective method of gaining vocabulary knowledge. However, teachers and students should understand that while it may be superior in developing receptive knowledge, productive learning might be better suited to improving speaking and writing skills. Taken as a whole the study indicates that there should be a role for both types of learning, and that role should be dependent on the type of learning that is required.

The second purpose of this study was to investigate the effects of receptive and productive learning tasks on the different aspects of knowledge. The results show that receptive and productive learning led to bi-directional gains in knowledge of all five aspects. This expands upon previous findings indicating that receptive learning may lead to gains in knowledge of orthography, syntax, association, grammar, and meaning (Experiments 1 and 2), and that receptive and productive learning from word pairs promotes bi-directional gains in meaning (Griffin and Harley, 1996; Stoddard, 1929; Waring, 1997b).

The results of Experiment 3 provide evidence that knowledge of association, syntax, and grammar may follow the same pattern of development as meaning. For each of these aspects, receptive learning led to higher mean scores on the receptive measures while productive learning led to higher scores on the productive measures. Orthography was the only aspect that did not follow this pattern. Productive learning led to significantly higher gains on both the receptive and productive measures. This result is interesting for two reasons. First it indicates that the relationships between the

aspects of knowledge may differ, and that differences may depend on how an item is learnt. Productive learning from word pairs focuses learners primarily on L2 form while receptive learning from word pairs focuses learners on L1 meaning. Not surprisingly, the highest combined score for the productive group was orthography while for the receptive group it was meaning. This result helps to show that different tasks are likely to have strengths and weaknesses pertaining to the types of knowledge that may be gained. For example, a task such as a word web in which items are learned in semantic sets is likely to promote large gains in paradigmatic and syntagmatic association, and meaning but lesser gains in grammatical functions and collocation. It is important that teachers consider what may or may not be gained through different activities. While this may seem obvious, until now there has been little data to support this. Earlier studies have focused primarily on meaning and ignored the other aspects of knowledge. Further research would be useful to help us understand how different tasks affect the different aspects.

The results of all three experiments showed that subjects made the lowest productive and receptive gains in knowledge of syntax. This was likely due to a lower degree of overlap between L1 and L2 syntax than association and grammar. Since the parts of speech of all of the target words were the same in the L1 and L2, the subjects could use their L1 knowledge to correctly answer the measures of grammatical functions. Similarly, they could use their L1 knowledge to produce or recognize words with a semantic relationship on the measures of association since paradigmatic associations are unlikely to change between L1 and L2. Rock, boulder, pebble, and stone have a semantic relationship regardless of the language. However, collocates often change between the L1 and L2 (Bahns, 1993) making it more difficult to use L1 knowledge to score correctly on L2 tests of syntax. For example, a common response to the target meaning *boulder* on the productive tests of syntax was *break*. In Japanese, *break* is a fairly common syntagmatic associate of boulder in the phrase 'kyoseki ga kudakeru'. In English, the two words may occur in context but it is not a common occurrence.

Taken as a whole, Experiments 3-5 indicate a strong relationship between the different aspects of knowledge. There was little difference between the mean scores for each group apart from syntax, which lagged slightly behind for the reasons detailed above. Since the tasks provided learners with both form and meaning it might be expected that those scores would be considerably higher. The size of the gains in these knowledge types in relation to association, grammar, and syntax indicates that the different aspects of knowledge may develop together. The range in mean scores on the productive measures for each task is shown in Table 35. The scores on the productive tests may provide a more accurate representation of the actual learning gains since they are identical in format and do not allow learners any strategies to score correctly through guessing. Mean scores on the receptive measures were also quite close but it is impossible to determine if any two multiple choice tests are equally sensitive since the distractors are different on each test.

Table 35 The range between the mean scores on the productive measures of orthography, association, meaning, and grammar.

| Experiment | Task | Range | Syntax |
|------------|---------------------------------|-----------|--------|
| 3 | Receptive word pairs | 7.15-7.76 | 4.82 |
| | Productive word pairs | 8.10-8.79 | 6.21 |
| 4 | Reading three glossed sentences | 8.71-9.43 | 7.29 |
| | Sentence production | 7.32-8.39 | 5.94 |
| | Reading three glossed sentences | 1.98-2.78 | 1.53 |
| 5 | Sentence production | 3.35-4.08 | 2.45 |

Table 35 shows that the size of the gains for the knowledge types was fairly consistent for each task. While it should not be surprising that reading and writing tasks promote knowledge of all five aspects of knowledge, it is surprising that this also occurs with receptive and productive word pairs. This result supports the findings of the previous two studies indicating that intermediate L2 learners may use L1 knowledge of the target words and/or L2 knowledge of their synonyms to gain knowledge of syntax, grammar, and association. Learners may initially use L1 meanings or L2 synonyms as models in order to produce or recognize paradigmatic and syntagmatic associates, and

the part of speech of L2 words. While this may lead to errors when they overgeneralize, it allows them the possibility of using the words. Gradually as learners gain L2 knowledge of the words through trial and error, and meetings in context they may build a more accurate representation of the different aspects.

(1996) suggest. If only one task were to be used, productive learning from word pairs may be better because of its superiority over the receptive task at promoting productive knowledge.

3. Productive learning from word pairs is better suited to gaining both receptive and productive knowledge of orthography. The productive task's focus on L2 form suggests that productive learning may be better suited to gaining knowledge of form. This result also indicates that different tasks may be more effective than others at promoting gains in certain aspects of knowledge.
4. Meeting target words in three glossed sentences and writing each target word in a sentence may be very effective methods of gaining knowledge of orthography, meaning, association, grammar, and syntax. When learners are given equal time to complete the tasks the receptive task may be more effective than the productive one. However, when time is limited to the amount needed to complete the task, writing may be superior to reading for both receptive and productive knowledge.
5. The similarity between the size of gains in knowledge of the different aspects suggests that intermediate learners may use their L1 knowledge of target words and/or L2 knowledge of synonyms to gain knowledge of syntax, grammar, and association. This supports the findings of Experiments 1 and 2.
6. Gains in syntax may be lower than association, meaning, grammar, and orthography due to a lesser degree of overlap between L1 and L2 syntax than with the other knowledge types.
7. Productive learning takes longer than receptive learning.

4.13 Rationale for the next experiment

The five previous experiments have investigated which aspects of vocabulary knowledge were gained through different learning tasks. A series of 10 tests has been used to isolate and measure gains in knowledge of orthography, meaning, association,

syntax, and grammar after learners have completed a task. The results have shown that all of the tasks contribute to gains in each aspect but one task in particular – learning from word pairs – has been surprisingly effective. The findings have indicated that the learning from word pairs task has led to large gains in knowledge of all aspects despite the fact that it is likely to take less time than all of the other tasks, and does not include context, which many researchers argue is necessary when learning vocabulary (Crow, 1986; Krashen, 1989).

While the different tests of vocabulary knowledge used in this thesis have provided a much more accurate assessment of what is gained from each task, the tests could be criticized because they do not measure communicative language gains. In other words, demonstrating that learners gain knowledge of the different aspects does not show that they can communicate with those words. Indeed, the majority of vocabulary tests can be criticized for lacking communicative value.

The final study in this thesis seeks to determine the size of communicative gains in knowledge through the learning from word pairs tasks. One test of receptive knowledge measuring comprehension, and another test of productive knowledge measuring writing will be used to assess communicative gains. This study may be useful for several reasons. First, very little research investigating the effects of pre-learning vocabulary on communicative knowledge has been completed. Since there is some debate about the merits of direct vocabulary teaching (Hulstijn, 2001; Nation, 2001; Tinkham, 1989), this study may provide evidence, which could help to determine the value of direct methods, and improve upon our understanding of L2 vocabulary acquisition. Second, the following study may help to show how knowledge of the different aspects affects communication. Third, determining whether learning from word pairs promotes communication skills would be helpful for teachers and learners. They could modify teaching and learning tasks to facilitate better communication skills.

CHAPTER 5

5.0 Introduction

The previous experiments in this thesis have shown that vocabulary instruction may lead to gains in receptive and productive knowledge of orthography, meaning, association, syntax, and grammar. In each experiment, the size of the vocabulary gains has provided insight into the effects of the different learning tasks. The results of Experiments 1, 2, and 3 have shown that learners make relatively large gains in vocabulary knowledge through learning from word pairs in comparison to the other tasks. This suggests that learning from word pairs is an effective method of learning vocabulary. However, this may be misleading because the relationship between vocabulary knowledge and communication is not clear. Past L1 research has shown that vocabulary instruction may improve vocabulary knowledge but may not affect comprehension (Jenkins, Pany, and Schreck, 1978; Pany and Jenkins, 1978; Tuinman and Brady, 1974) or writing (Duin and Graves, 1987). Moreover, L1 studies that have investigated the effects of different vocabulary learning tasks tend to show that intensive learning tasks that involve both contextual and definitional learning may facilitate comprehension (Beck, Perfetti, and McKeown, 1982; McKeown, Beck, Omanson, and Perfetti, 1983) and writing (Duin and Graves, 1986, 1987), while decontextualized tasks that focus on meaning such as learning from word pairs might be ineffective (Mezynski, 1983; Stahl and Fairbanks, 1986). However, the results have been inconsistent, and there is some evidence that decontextualized tasks may also be effective (McDaniel and Pressley, 1984, 1989; Pressley, Levin, and Miller, 1981).

Surprisingly, no L1 and L2 studies have compared the effects of receptive and productive learning on comprehension and writing. One of the more interesting findings in Experiment 3 of this thesis was that productive learning from word pairs contributed to significantly higher scores on all of the productive tests and on the receptive test of orthography, while the receptive learning from word pairs task was

better on the receptive tests of association, grammar, syntax, and meaning, although the only significant difference was for meaning. The results were supported by previous research indicating that productive learning is more effective in increasing productive knowledge, and receptive learning is more effective in increasing receptive knowledge (Griffin and Harley, 1996; Stoddard, 1929; Waring, 1997b). If receptive and productive learning have a similar effect on communication, receptive learning may be better suited to improving receptive tasks such as comprehension, and productive learning may be better suited to improving productive tasks such as writing.

The contradictory L1 findings and the lack of L2 studies indicate that further research is warranted. The present study investigated the effects of receptive and productive learning from word pairs on comprehension, and the use of taught words in writing. Specifically it examined the following questions:

1. Can direct vocabulary learning result in successful understanding of the vocabulary in reading?
2. Does receptive learning from word pairs result in more successful comprehension of the vocabulary in reading than productive learning from word pairs?
3. Can direct vocabulary learning result in successful use of the vocabulary in writing?
4. Does productive learning from word pairs result in more successful use of the vocabulary in writing than receptive learning from word pairs?

5.1 How does vocabulary knowledge affect comprehension?

L1 research investigating the effects of vocabulary instruction on comprehension has been inconsistent. Some studies have found that vocabulary instruction has a significant effect on vocabulary knowledge but may not affect comprehension (Jenkins, Pany, and Schreck, 1978; Pany and Jenkins, 1978; Tuinman and Brady, 1974) while others have shown that it may improve both vocabulary knowledge and comprehension

(Beck, Perfetti, and McKeown, 1982; Kameenui, Carnine, and Freschi, 1982; McDaniel and Pressley, 1989; McKeown, Beck, Omanson, and Pople, 1985). The majority of studies have found that rich instruction has been most effective (Beck, Perfetti, and McKeown, 1982; Kameenui, Carnine, and Freschi, 1982; McKeown, Beck, Omanson, and Perfetti, 1983), and methods that emphasize linking meaning with form have been least effective (Mezynski, 1983, Stahl and Fairbanks, 1986). However, these findings are also inconsistent, and there is some evidence indicating that decontextualized learning tasks such as the keyword technique and learning from word pairs may also increase comprehension (McDaniel and Pressley, 1989).

In a review of eight L1 studies that investigated the effects of vocabulary knowledge on comprehension, Mezynski (1983) reports that there are three characteristics of vocabulary instruction that succeeded in improving comprehension: increasing the amount of time spent on tasks, using a variety of tasks to develop a deeper knowledge of target words, and encouraging productive learning. She suggests that decontextualized learning tasks such as word pairs are less likely to facilitate comprehension than multiple context learning. In a meta-analysis of 52 studies that investigated the relationship between vocabulary and comprehension, Stahl and Fairbanks (1986) reached a similar conclusion. They report that the methods that had the strongest effects involved both definitional and contextual information, and that instruction which provided only definitional information about target words such as paired associate learning did not have reliable effects on comprehension. However, the superiority of rich instruction may be slightly misleading because in at least two of the studies (Beck, Perfetti, and McKeown, 1982; McKeown, Beck, Omanson, and Perfetti, 1983), the time spent on the tasks was likely to be much more than is normally allotted for vocabulary instruction (Beck, McKeown, and Omanson, 1987).

Perhaps it should not be that surprising that rich instruction facilitates comprehension. If vocabulary knowledge does affect comprehension as most researchers suggest, intensive tasks are most likely to be effective. What might be more useful, however, would be to determine whether commonly used tasks that are relatively time efficient

are effective. McDaniel and Pressley's (1989) study provides some evidence that learning from word pairs and the keyword technique may improve comprehension.

McDaniel and Pressley (1989) compared the effects of learning from the keyword method, learning from word pairs, and a combined learning from context and learning from word pairs task on vocabulary knowledge and comprehension in two experiments. After the treatments, the subjects read a short passage in either an embellished or unembellished format. The embellished format provided contextual clues to help point out the meaning of the 15 L1 target words. Vocabulary knowledge was measured by a receptive recall test in which learners had to write the definitions of the target words. Comprehension was measured by a true/false test, a cued recall test, and a retell test (free recall). Questions in the true/false test referred to phrases in the passage that contained target items. In the retell test, the subjects had to recall as much of the passage as they could, and in the cued recall test, the subjects had to answer questions about the information that the target words conveyed in the passage. The results showed that the keyword method produced higher levels of receptive recall (76%) than the combined (53%), and word pairs methods (52%) on the test of vocabulary knowledge. Mean percentage scores on the true/false test were also significantly higher for the keyword group (95%), with little difference between the scores for the combined context and meaning group (89%), and the word pairs group (87%). Vocabulary instruction did not affect free recall on the retell test; however, the mean score for the word pairs group was significantly higher than a no-exposure control group. On the test of cued recall, subjects that completed the learning from word pairs, and the keyword tasks scored significantly higher than the combined method.

McDaniel and Pressley's study is valuable because it shows that decontextualized learning tasks such as learning from word pairs and the keyword method may improve comprehension, and they are not necessarily inferior to tasks that involve both definitional learning and contextual learning. This contrasts with previous studies reviewed by Stahl and Fairbanks (1986) and Mezynski (1983).

While there have been many L1 studies investigating the effects of vocabulary instruction on comprehension, there have been very few L2 studies. The only L2 study that I am aware of was conducted by Johnson (1982). She found that studying the definitions of target words prior to reading a passage had no significant effects on two comprehension tests. However, very little detail was given about the instruction including how much time was spent on the task.

The following section attempts to provide an explanation for the inconsistency in the findings.

5.2 Factors that may affect comprehension

Research investigating the effects of rich instruction on comprehension indicates that vocabulary knowledge is a factor that affects reading comprehension. While this should not be a surprise, it is puzzling that few other types of vocabulary instruction have had much effect. One reason for the lack of significant findings is that there may be many factors involved in understanding a text. As Stratton and Nacke (1974: 185) point out, "reading is a complex information-processing task and identifying word meanings is but one element in the process." Although research investigating the effects of vocabulary instruction on comprehension has been inconclusive, it has been very successful at identifying and demonstrating that many elements or factors are involved in understanding a text.

Several studies have shown that prior knowledge about a topic or background knowledge affects comprehension. Stahl and Jacobson (1986), and Stahl, Jacobson, Davis, and Davis (1989) found that pre-teaching a topic can significantly improve understanding of a text about an unfamiliar issue, and Stahl, Hare, Sinatra, and Gregory (1991) found that background knowledge had a significant effect on what was understood, as well as how much could be recalled about a text. The studies indicate that the more learners know about an issue, the more likely they are to understand a

text about that topic. Learners may know all of the words in a text but if they have little knowledge of that topic, it may affect their performance on a comprehension test.

Individual differences may also affect comprehension (Mezynski, 1983; Stahl, 1990). Students who know more words, are more proficient in the language, and have a better memory, are more likely to understand a text, and perform well on a comprehension test. A third factor that may affect the understanding of a text is context (Mezynski, 1983; Stahl, 1990). Stahl (1990) reports that three contextual factors that affect comprehension are: (a) the overlap between the meaning of the context and the word's meaning, (b) the amount of information available in the context, and (c) the importance of the word in the passage. Mezynski (1983) reports that redundant information in the context, and syntactic cues may provide enough information for learners to correctly answer comprehension questions. Kameenui, Carnine, and Freschi (1982) found that comprehension scores were higher for texts containing redundant information that allowed the readers to make inferences about the topic of the passage and difficult vocabulary items in both experiments. If the meanings of target words can be inferred through the context, vocabulary instruction may have little affect on comprehension.

A fifth factor is the density of unknown words in a passage (Mezynski, 1983; Pany and Jenkins, 1978). Vocabulary instruction may have little effect if there is still a large number of unknown words in the text. Similarly, Stahl and Fairbanks (1986) suggest that a text must contain a large proportion of taught words for vocabulary instruction to affect comprehension. Teaching a small number of words in a text is less likely to affect comprehension than teaching a large number of words. Moreover, if unknown words convey important information, a passage may not be clearly understood. Wixson (1986) found that instruction on words that were important to the central ideas of a text facilitated comprehension of those story lines while instruction on words that were of less importance to the text increased comprehension of noncentral story lines.

There are several vocabulary factors that may also affect comprehension. Baker (1989) suggests that for less experienced learners the length of unknown words may affect

comprehension. Longer words are more likely to disrupt understanding of a passage than shorter words. Stahl (1990) reports that conceptual complexity, and polysemy and homonymy may affect comprehension. It is also likely that some of the interlexical and intralexical factors that affect L2 vocabulary learning may also affect L2 comprehension (A list of these factors is given in Section 3.4.14). However, it should be noted that since most research involves learning the same words in each task, vocabulary factors were unlikely to have affected the results in any of the studies.

Perhaps the most important factor in determining the effects of vocabulary instruction is the type of tests used to measure comprehension. Several different types of tests have been administered such as retell tests or tests of free recall (Kameenui, Carnine, and Freschi, 1982; McDaniel and Pressley, 1989; McKeown, Beck, Omanson, and Pople, 1985; Pany, Jenkins, and Schreck, 1982), open ended comprehension questions (McDaniel and Pressley, 1989; Pany and Jenkins, 1978), multiple-choice comprehension questions (Kameenui, Carnine, and Freschi, 1982; McKeown, Beck, Omanson, and Pople, 1985), true/false questions (McDaniel and Pressley, 1989; Pany, Jenkins, and Schreck, 1982; Pressley, Levin, and Miller, 1981), and a paraphrase test (Pany, Jenkins, and Schreck, 1982). There are three reasons why tests may be a factor. First, tests designed to measure understanding of story lines that incorporate the instructed vocabulary may be more sensitive to comprehension than recall tests (Wixson, 1986). Second, since results often vary between comprehension tests, different test formats may tap different types of knowledge (McDaniel and Pressley, 1989), and third, the vocabulary used in comprehension questions may affect students' performance (Cunningham and Moore, 1993). If the type of tests used to measure comprehension affect performance then researchers must be very careful when selecting a test format. Some tests may show that students understood a text when in fact they did not, and others may show that they did not understand when they did.

Taken as a whole, the research indicates that many factors may affect comprehension. Controlling all of the factors is essential to determine the effects of vocabulary instruction. If some of the factors are not controlled, the effects of vocabulary

knowledge on comprehension may be reduced or even eliminated. The present study examined comprehension of sentences rather than text to reduce the number of factors that may affect the results. Since there has been virtually no L2 research investigating the effects of vocabulary knowledge on comprehension, positive results would indicate that more research is warranted.

While the research has been very useful, two questions remain unanswered. First, does L2 vocabulary instruction facilitate L2 comprehension? At present, there does not appear to be any evidence to support this. Experiment 2 from this thesis indicates that learning from word pairs is a very effective method of gaining L2 vocabulary knowledge, and may be as effective as methods that provide definitional and contextual learning. However, the L1 studies demonstrate that while vocabulary instruction has consistently had a significant effect on vocabulary knowledge, it does not always affect comprehension. The present study seeks to shed some light on this issue. Second, do receptive and productive tasks have a different effect on comprehension? Perhaps because of the inconsistent findings, there has not been any research comparing the effects of receptive and productive learning tasks. Mezynski (1983) suggests that productive tasks may be more effective in improving comprehension, and previous research comparing receptive and productive learning from words pairs suggests that productive learning is more effective in increasing productive knowledge of meaning while receptive learning is more effective in increasing receptive knowledge of meaning. Experiment 3 supported the earlier findings, and also indicated that the productive task may be more effective in increasing productive knowledge of association, syntax, grammar, and orthography, and receptive knowledge of orthography, and the receptive task may be more effective in increasing receptive knowledge of meaning. However, since reading comprehension is essentially a receptive task, the findings would suggest that receptive learning from word pairs may be more effective in improving comprehension. One aim of this study is to determine the effects of receptive and productive word pairs on comprehension.

5.3 How does vocabulary knowledge affect writing?

At present, there is very little research that has investigated the effects of vocabulary knowledge on writing. L1 research tends to show that learners may be able to successfully use recently taught words in their writing, and that their performance might depend on the type of instruction (Duin and Graves, 1987; McDaniel and Pressley, 1984; Pressley, Levin, and Miller, 1981). In a review of the research, Graves (1986: 62-63) reports that "it does seem likely that vocabulary instruction can improve some aspects of students' writing and that instruction intended to increase students' use of the words taught should directly encourage them to use the words."

Duin and Graves (1986) found that rich instruction led to significant gains in vocabulary knowledge, and improved the quality of writing in narratives in comparison to a no-instruction control group. The results also showed that fourth grade students that received instruction used 80 percent of the words in their narratives, and sixth graders used 40 percent. However, the fourth graders did not always use the words accurately.

A second study by Duin and Graves (1987) examined the effects of three instructional procedures on vocabulary knowledge, and writing. Each teaching method combined receptive and productive tasks but varied the amount of writing students did. The results indicated that for the learning conditions that involved the most writing, students had significantly higher scores on the vocabulary measure, used more of the words in essays, and had superior writing to the group that did the least writing while learning. The results also showed that learners may gain vocabulary knowledge but may not improve their writing. One group was able to demonstrate knowledge of 75% of the target words on the vocabulary test but used an average of only 0.03 of the words in their essays, and ratings of their writing performance decreased on their post-test essays. This suggests that vocabulary tests alone may not provide an accurate indication of writing ability. While tests of vocabulary knowledge might provide

useful data, sentence production tests may give a more precise measure of what learners can or cannot do with a word.

Pressley, Levin, and Miller (1981) compared the effects of the keyword method and learning from word pairs on writing. At testing the subjects were provided with a list of the 32 target words, and were instructed to write each item in a sentence in which the meaning of the target words could be derived. They found that university students who had learned L1 words with the keyword technique had higher scores on a cloze test, and were able to write significantly more correct sentences (47.5%) than students who had been learning from word pairs (32.8%).

McDaniel and Pressley (1984) compared incidental learning from three sentences and the keyword method. After the treatments, the subjects were given a receptive test of recall, and were required to write 11 of the 61 target words in two sentences. The results showed that the subjects that learned with the keyword technique could recall significantly more of the target words, and could generate more correct sentences than the incidental learning group. McDaniel and Pressley (1984) also found that performance on the sentence production task was related to recall of meaning. If a word's meaning was recalled, the proportion of those words that were correctly written in a sentence was .91 in the keyword condition and .85 in the context condition.

Taken as a whole, the results indicate that L1 learners may be able to use pre-taught words in their writing, with performance dependent on the method of learning. The findings also suggest that vocabulary knowledge may also be an indicator of writing performance. In all of the studies, one learning condition was significantly more effective on the vocabulary measures and the writing tests. However, this relationship needs to be investigated further because in Duin and Graves' (1987) study one group of subjects who demonstrated large gains in vocabulary knowledge did poorly on writing tests. Duin and Graves' (1986, 1987) studies suggest that productive learning tasks may be effective, and the studies by Pressley, Levin, and Miller (1981) and McDaniel and Pressley (1984) indicate that gaining knowledge of meaning may be critical.

Since there have been so few studies, further investigation of the effect of vocabulary instruction on writing are needed. This is particularly apparent with L2 research for which there do not appear to be any studies that have specifically addressed this issue. Furthermore, no studies have compared the effects of receptive and productive learning on writing. The results of previous research comparing receptive and productive learning from word pairs (Griffin and Harley, 1996; Stoddard, 1929; Waring, 1997b), and Experiment 3 indicate that the productive task is more effective than the receptive task in increasing productive knowledge. Since writing is essentially a productive task, the findings suggest that productive learning from word pairs may be more effective.

The present study examined the relationship between L2 vocabulary learning and writing. Specifically, it investigated the following two questions:

1. Can direct vocabulary learning result in successful use of the vocabulary in writing?
2. Does productive learning from word pairs result in more successful use of the vocabulary in writing than receptive learning from word pairs?

5.4 Implementation of the research

5.4.1 Research questions

The present study will address the following questions:

1. Can direct vocabulary learning result in successful understanding of the vocabulary in reading?
2. Is receptive learning from word pairs more effective than productive learning from word pairs in improving comprehension as learning from word pairs research suggests or less effective as Mezynski (1983) suggests?

3. Can direct vocabulary learning result in successful use of the vocabulary in writing?
4. Is the productive word pairs task more effective than the receptive task in improving the use of instructed words as Duin and Graves' (1986, 1987) studies and learning from word pairs research suggest?

5.4.2 Pilot studies

Pilot studies conducted with Japanese learners studying English as a foreign language determined the following:

1. The subjects were unlikely to have prior L2 knowledge of the 15 target words.
2. The subjects could complete the treatments in six minutes, and the dependent measures in 80 minutes, and it was unlikely that there would be ceiling or floor effects on the tests. If the subjects were given more time to complete the tasks, the subjects may perform better on the dependent measures but there was the possibility of a ceiling effect on the receptive vocabulary tests.
3. The productive task was likely to be more effective than the receptive task on the picture description test.
4. The subjects would be able to understand the contexts on the comprehension test.
5. The subjects should be able to answer the picture description and comprehension tests correctly if they know the target words.

As a result of the pilot studies, the following changes were made to the study:

1. Time was limited to six minutes in the treatments.

2. Target words, visual prompts on the picture description test, and questions on the comprehension test were selected.

5.4.3 Subjects

The subjects in this experiment were 71 Japanese native speakers from three first-year EFL classes at Kyushu University in Fukuoka, Japan. All of the subjects had studied English for a minimum of six years, and scored 80 percent or higher on Version 1 of the Vocabulary Levels Test at the 2nd 1000 level. Their mean score was 27.3 indicating that they were well in control of that level, and that they had receptive knowledge of almost all of the 2000 most frequent words (Schmitt, Schmitt, and Clapham, 2001). Their mean score on version C of the Productive Vocabulary Levels Test at the 2nd 1000 word level was 14.2 suggesting that they had productive knowledge of approximately 1700 of the 2000 most frequent words. The subjects were randomly assigned to the experimental groups.

5.4.4 Design

Two experimental groups were used in this study to examine the effects of receptive and productive learning of word pairs on comprehension and writing. One experimental group studied 15 L2 target words receptively, and the other studied the same target words productively. In the receptive treatment, the target words were presented in a column on the left side of a paper, and their translations were presented on their right. The subjects were instructed to cover the translations and then look at the target words and try to recall their translations. If they could not recall a meaning, the subjects were told to uncover the translation and check the meaning. In the productive treatment, the translations were on the left and the target words were on the right. The subjects completed the same task except they covered the target words, and then tried to recall them. Each group was given six minutes to complete their task.

Twelve tests measuring writing, comprehension, and vocabulary knowledge were administered after the treatments. The writing test was given first followed by the comprehension test, and then 10 tests of vocabulary knowledge. The vocabulary tests used the same format as those used in Experiments 2-5, and the subjects completed the same procedure used in the previous experiments when completing the tests. The results of the groups were then compared to determine how receptive and productive learning from word pairs affects writing, comprehension, and vocabulary knowledge.

5.4.5 Procedure

The experiment was conducted in three stages over two weeks. The first stage of the experiment was mandatory for all students in the classes while the second and third stages were optional. In the first stage, students completed the receptive and productive Vocabulary Levels Tests at the 2nd 1000 word level. In the second stage of the experiment, two subjects who did not take part in stage 3 were involved in case studies outside of regular class time. These subjects were randomly assigned to the experimental groups. While they completed the treatment and dependent measures they were asked to think out loud to provide insight into subjects' thought processes during the experiment. Their results were not included in stage three. During the third stage of the experiment, the students completed the treatments and tests.

The experimental procedure was implemented over the following three stages:

- | | |
|---------|--|
| Stage 1 | The Vocabulary Levels Tests were administered to determine the vocabulary size of potential subjects. Subjects were randomly assigned to treatment groups. |
| Stage 2 | Case studies were conducted. |
| Stage 3 | The subjects participated in treatments, and the dependent measures were administered. |

At the beginning of stage 3, the instructor explained and demonstrated the two tasks. The subjects were told that they would have six minutes to learn the 15 target words, and they would be tested after the treatment. Any questions were addressed to make certain that the instructions were completely understood. The subjects were not given any information about how they would be tested. Each task was presented on one page with instructions written in Japanese at the top of the page. During the treatment, the instructor monitored the subjects to ensure that they were using the desired learning tasks.

After the treatments, the subjects were given the writing test. When all of the subjects had finished, it was collected and the comprehension test was administered. After the comprehension test was completed, it was collected and the subjects were given the 10 page vocabulary test booklets. There was one test on each page, and the subjects were instructed to complete them in order. Once a test was finished they were not allowed to go back to it and revise any answers. They were carefully supervised to ensure that this did not occur. This was to eliminate the possibility of a learning effect from subsequent tests. Subjects had as much time as they needed to finish a test.

5.4.6 Materials

5.4.7 Target words

The following 15 target words were used in this experiment: boulder, constellation, dagger, doze, faucet, foal, kneel, lane, lick, locomotive, reef, sob, spear, sprint, and squirt. The target words were comprised of nine nouns and six verbs. The ratio of nine nouns to six verbs was used because noun and verbs are the most common parts of speech found in natural text, and the 9 : 6 ratio approximates their proportional frequency of occurrence (Kucera and Francis, 1967). All of the target words were taken from the fifth frequency band in the COBUILD dictionary. It accounts for the 6601st to 14,700th most frequent words. The target words were replaced with 15 nonsense words created and used in Experiments 1 and 2 to ensure that the subjects had

no prior knowledge of the target words. All of the nonsense words were two syllables. Eleven of them were five letters long, and four of them were six letters. The nonsense words were as follows: ancon, cader, copac, dangy, dapew, denent, faddam, gishom, hodet, masco, mesut, nasin, pacon, sagod, and tasper.

5.4.8 Word Pairs

In the treatment, the fifteen nonsense words were presented together with their L1 translations on one page. In the receptive task, the nonsense words were on the left of the translations (see Appendix L). In the following example, the target word *locomotive* was replaced with the nonsense word *masco*, and paired with the Japanese translation of locomotive,

masco 機関車

whereas, in this example taken from the productive task, the L1 translation is on the left of the target word (see Appendix M).

機関車 **masco**

5.4.9 Time

Subjects were given six minutes to complete the treatments. It was determined in pilot studies that six minutes was sufficient for the subjects to finish their tasks.

5.4.10 Dependent Measures

The first test that was administered after the treatments was a picture description test (see Appendix N). It was used to measure the subjects' use of the target words in writing. In this test, the subjects were presented with 15 pictures, and asked to write

about each one using the target words. The criteria for selecting the pictures were (a) there was an obvious visual relationship with a target word, (b) subjects in pilot studies could identify the target word that was depicted in each picture, and (c) subjects in pilot studies could write correct sentences using the target words for each picture. The subjects were given the following instructions in English and in the subjects L1: "Write one sentence about each picture using a word that you learned today." The following examples are for the target words *lick*, *dagger*, and *spear*:



1. _____ 2. _____ 3. _____

To score correctly, the subjects had to write a correct sentence about each picture using the target word. Since the test was designed to measure communicative knowledge rather than vocabulary knowledge small mistakes in spelling were acceptable. For example, for the target word *denent* the following spellings: *denant*, *denen*, and *donent*, would all be acceptable because each response is very close to the correct spelling, and they do not resemble any of the other target words. However, *denew*, *denut*, and *dapent*, would be unacceptable spellings because each response partially resembles another target word.

The 'picture description' test format was chosen for four reasons. First, it was relatively quick and clear in comparison to some other test formats such as essay writing. Second, students needed to use the target words to score correctly in the task. In the free production test format used in Duin and Graves (1986, 1987), students are

not required to use the target words. Thus, in that format it is uncertain whether the subjects are unable to use the target words or simply do not feel it is necessary to use them. Third, the test format allowed unrelated target words to be used. Tinkham (1993) and Waring (1997a) have shown that learning semantically related words together could reduce vocabulary gains. Typically, in free production, learners write an essay about a topic that involves all of the target words. Therefore, since the semantic relationship between the target words may hinder learning, the effect of the learning task would not be clear. Fourth, the picture description format ensures that students will demonstrate semantic knowledge of the target words but in sentence production tests this does not always occur. Students could write sentences such as 'It is a dagger.' and 'There is a dagger.' in sentence production tests that only indicate grammatical knowledge of the target word *dagger*.

The comprehension test followed the picture description test. The comprehension test used a true/false format. Each of the 15 target words was presented in two sentences making a total of 30 sentences in the test (see Appendix O). Each sentence was constructed so that the target words were critical for accurate comprehension, and their meanings could not be derived from the running words in the sentences. Comprehension of a sentence was necessary to score correctly. The following examples are for the target words *sob* (pacon) and *spear* (denent):

| | | |
|---|------|-------|
| People often pacon when they are sad. | True | False |
| Men never pacon. | True | False |
| Students sometimes bring denents to school. | True | False |
| Denents are very heavy. | True | False |

The true/false test format was chosen over several other formats such as comprehension questions and free recall. The true/false format in this experiment measured sentence comprehension rather than overall comprehension of a text. Tests that measure textual comprehension such as free recall and comprehension questions may be the least sensitive of the test types because memory rather than vocabulary instruction might be the critical factor in scoring correctly. Rystrom (1970) suggests that memory is the best

single predictor of reading comprehension. Thus, learners may understand a text but if they do not concentrate on remembering it, their performance on a test may suggest that their comprehension was poor. Similarly, learners may not understand the details in a passage but if they can remember them, their performance on the tests may indicate that they had excellent comprehension. A second problem with free recall and open ended comprehension questions is that while learners may understand a text, they may not be able to describe it. Comprehension is a receptive process; however, those test formats require learners to demonstrate productive knowledge. As was mentioned in Experiment 3, most researchers suggest that learners are likely to have a larger receptive vocabulary than productive vocabulary. Therefore, there is the possibility that a learner may perform poorly on a test because they lack productive vocabulary knowledge.

The vocabulary tests followed the comprehension test. The same ten tests that were administered in Experiment 2 were used to measure the effects of receptive and productive learning tasks on vocabulary knowledge in Experiment 6 (for an example of the test formats see Appendix F). Each test enabled learners to demonstrate a specific aspect of word knowledge productively or receptively. The tests measured receptive and productive knowledge of orthography, meaning and form, grammatical functions, association, and syntax.

5.5 Case studies

Case studies were done with two learners. Each subject completed one of the treatments. In the case studies, the subjects were asked to think aloud during both the treatment and tests. After the completion of the tests, questions were asked to gain further insight into the learning process.

The following observations emerged from the case studies:

1. Both subjects found the picture description test more difficult than the

comprehension test.

2. The critical factor in the writing test appeared to be linking form and meaning. When the subjects were cued with the target words on questions they could not answer, they were usually able to respond correctly.
3. When asked to provide more information about their responses in the writing test, the subjects had little difficulty in doing so.
4. The subjects were able to successfully answer most of the questions on the comprehension test.
5. Despite having gained knowledge of meaning and form, some comprehension questions proved difficult. One subject responded incorrectly on the following two questions despite knowing that the meaning of the target word *ancon* was *dagger* and *denent* was *spear*: *Ancons are made of wood. Denents are very heavy*. This appeared to be due to poor understanding of those target words in the subjects' L1.
6. The subjects indicated that the primary difficulty on the productive tests of meaning, association, syntax, and grammar was recalling meaning and form.
7. Both subjects looked for hints in the distractors in the receptive tests of grammar, syntax, and association to help them respond correctly.

5.6 Results

The descriptive statistics (means, standard deviations, and number of subjects) of the scores for the communicative and vocabulary knowledge tests are reported in Tables 36 and 38. To determine whether there were any overall differences among the treatment groups, a multivariate analysis of variance (MANOVA) was performed using the scores on the 12 dependent measures (picture description, comprehension, productive orthography, receptive orthography, productive meaning and form, receptive meaning and form, productive association, receptive association, productive syntax, receptive syntax, productive grammar, and receptive grammar). The independent variable was the type of learning task (receptive and productive word pairs). The MANOVA revealed an overall statistically significant difference (Pillai's Trace, Hotelling's Trace,

Wilks' Lambda, and Roy's Greatest Root) between the two tasks ($F(12,58) = 8.52$, $p < .000$).

Table 36 Means and standard deviations of picture description, and comprehension tests for the two learning conditions

| Learning Condition | <i>n</i> | Picture description | | Comprehension (out of 30) | | Comprehension (out of 15) |
|-----------------------|----------|---------------------|--------|---------------------------|--------|---------------------------|
| Productive word pairs | 33 | 6.24 | (3.14) | 22.82 | (3.51) | 9.55 (2.03) |
| Receptive word pairs | 38 | 4.39 | (2.78) | 25.08 | (3.10) | 10.89 (2.29) |

Note. standard deviations are in parentheses; Comprehension (out of 15) required both responses for each target word to be correct.

Table 37 Summary of multivariate analysis of variance on picture description and comprehension tests

| Picture description | Comprehension (out of 30) | Comprehension (out of 15) |
|---------------------------------|------------------------------------|---------------------------------|
| $F(1,69) = 6.92$, $p = .010^*$ | $F(1,69) = 8.31$, $p = .005^{**}$ | $F(1,69) = 6.81$, $p = .011^*$ |

Note. $^*p < .05$, $^{**}p < .01$

In answer to research questions 1 and 3, the results suggest that L2 vocabulary learning may improve both comprehension and writing. Table 36 shows that the productive learners had a mean score of 6.24 on the picture description test, and the receptive learners had mean score of 4.39. On the comprehension test, the receptive group had a mean score of 25.08 and when the criteria for a correct response was that both answers for each target word were correct, they had a mean score of 10.89. The productive group had mean scores of 22.82 on the comprehension test, and 9.55 when it was scored out of 15. The correlation between the two conditions for a correct response on the comprehension test was found to be .933 ($p < .000$). This indicates that it may not be necessary to use two correct answers for one word as the criterion to score correctly on a true/false comprehension test.

In answer to research question 2, the results indicate that the receptive learning from word pairs group had significantly higher scores on the comprehension test ($F(1,69) =$

8.31, $p < .01$) indicating that receptive learning from word pairs is more effective in improving comprehension. The results also show that when both responses for each target word were correct, the scores on the comprehension test were also higher for the receptive group ($F(1,69) = 6.81$, $p < .05$). In answer to research question 4, the MANOVA revealed that the subjects that completed the productive learning from word pairs task had significantly higher scores on the picture description test ($F(1,69) = 6.92$, $p < .05$) indicating that productive learning from word cards is more effective in improving the use of students' taught words in writing. A summary of the findings is shown in Table 37.

Table 38 Means and standard deviations of vocabulary tests for the two learning conditions in Experiment 6

| Learning Condition | <i>n</i> | PO | RO | PM | RM | PA | RA | PS | RS | PG | RG |
|-----------------------|----------|-----------------|-----------------|----------------|----------------|----------------|-----------------|----------------|-----------------|-----------------|-----------------|
| Productive word pairs | 33 | 11.79 (2.53) | 13.24 (2.17) | 9.64 (3.40) | 9.39 (3.54) | 8.21 (3.28) | 10.88 (2.48) | 7.55 (3.97) | 10.91 (2.94) | 10.03 (3.75) | 12.64 (2.82) |
| Receptive word pairs | 38 | 8.37 (2.96) | 11.82 (1.97) | 7.66 (4.03) | 9.95 (3.44) | 7.45 (3.22) | 11.74 (2.30) | 6.37 (3.11) | 11.34 (2.86) | 9.11 (3.68) | 12.74 (2.36) |

Note. Maximum score = 15, standard deviations are in parentheses, Key. PO=productive knowledge of orthography, RO=receptive knowledge of orthography, PM=productive knowledge of meaning and form, RM=receptive knowledge of meaning and form, PA=productive knowledge of association, RA=receptive knowledge of association, PS=productive knowledge of syntax, RS=receptive knowledge of syntax, PG=productive knowledge of grammar, RG=receptive knowledge of grammar.

Table 39 Summary of multivariate analysis of variance on vocabulary tests

| | productive | receptive |
|-------------|--------------------------------------|------------------------------------|
| orthography | $F(1,69) = 26.87$, $p = .000^{***}$ | $F(1,69) = 8.44$, $p = .005^{**}$ |
| meaning | $F(1,69) = 4.92$, $p = .030^*$ | $F(1,69) = 0.45$, $p = .507$ |
| association | $F(1,69) = 0.98$, $p = .326$ | $F(1,69) = 2.28$, $p = .135$ |
| syntax | $F(1,69) = 1.96$, $p = .166$ | $F(1,69) = 0.39$, $p = .532$ |
| grammar | $F(1,69) = 1.10$, $p = .299$ | $F(1,69) = 0.03$, $p = .870$ |

Note. $^*p < .05$, $^{**}p < .01$, $^{***}p < .000$

Table 38 shows that the productive task was more effective on the productive test of orthography ($F(1,69) = 26.87$, $p < .000$), the receptive test of orthography ($F(1,69) = 8.44$,

$p < .01$), and the productive test of meaning ($F(1,69) = 4.92, p < .05$). The productive learning group also outperformed the receptive group on the productive tests of association, syntax, and grammar but the differences were not significant. In contrast, the receptive group demonstrated larger gains on all of the receptive measures except for orthography; however, none of the gains were significant. A summary of the results of the MANOVA is shown in Table 39.

5.7 Discussion

The first question investigated by this research focused on the effect of pre-learning vocabulary on comprehension. The results indicate that L2 students who have studied target vocabulary may be able to understand sentences containing those words. The subjects demonstrated that they understood 80 percent of the sentences on the comprehension test, and 68 percent when the criterion for a correct response was both sentences correct for each target word. This result is supported by L1 research that has shown that vocabulary instruction improves comprehension (Beck, Perfetti, and McKeown, 1982; Kameenui, Carnine, and Freschi, 1982; McDaniel and Pressley, 1989; McKeown, Beck, Omanson, and Pople, 1985), and also lends support to McDaniel and Pressley's (1989) findings which showed that decontextualized tasks might facilitate comprehension. Since Experiment 6 was one of only two L2 studies that were designed to investigate the effects of vocabulary learning on comprehension, and the first to show that vocabulary knowledge increases comprehension, more research is warranted. In particular, the relationship between comprehension and test format needs to be examined. The true/false test administered in this study must be considered a somewhat narrow measure of comprehension because it measured sentence comprehension rather than text comprehension. Since understanding a sentence tends to be easier than understanding a passage the results may be different using another experimental design. Further research needs to explore how different tasks and tests affect comprehension. Perhaps, investigating comprehension using several tests may not only show that learners understand a text, but that there may also

be a depth of understanding similar to the way multiple vocabulary tests have found that there is a depth of vocabulary knowledge.

The second question examined by this research investigated the relative effectiveness of receptive and productive learning from word pairs on comprehension. A comparison of the two tasks indicated that receptive learning contributed to significantly higher scores on the comprehension test than productive learning. The receptive and productive groups had mean percentage scores of 84 and 76 on the comprehension test, and when the criterion for a correct response was both sentences correct for each target word, their scores were 73 and 64, respectively. The results indicate that receptive vocabulary learning may be more beneficial to understanding a text than productive learning. This is supported by earlier findings that have shown that receptive learning from word pairs is better suited to developing receptive vocabulary knowledge (Experiment 3, Griffin and Harley, 1996; Stoddard, 1929; Waring, 1997b). Therefore, if the primary aim of instruction is to improve comprehension, receptive tasks may be more effective.

Since the students that completed the receptive task did significantly better on the comprehension test, it might be expected that they would also do better on the receptive vocabulary tests. However, this did not occur. Although the receptive group outperformed the productive group on the receptive tests of meaning, association, syntax, and grammar, none of the differences were significant. There are three possible explanations for the lack of significant findings. First, the multiple-choice format may have affected the results. In the case studies from this experiment and several of the other experiments, subjects have repeatedly attempted to learn from the distractors on the receptive tests. Perhaps, this could have reduced the difference in gains on those tests. A second reason is that the comprehension test may have contributed to a learning effect. Since the comprehension test was essentially a receptive task, it may have aided the productive learners more than the receptive learners. This may account for the fact that this is the first study that did not find that the receptive learning from word pairs group had significantly higher scores on the receptive recall test. The third

explanation is that the receptive task is not significantly more effective at increasing receptive knowledge. Since the results of Experiment 3 also found no significant differences on the receptive tests of syntax, association, and grammar, there may not be as much difference between the two tasks in gaining receptive knowledge. However, since the experiments in this thesis are the first to compare gains for several aspects of knowledge, further research is needed.

In a comparison between the comprehension test and the vocabulary tests, there appears to be a strong relationship between receptive recall and comprehension; however, the link between the two was probably distorted by the experimental design. The proportion of correct responses on the receptive recall test, given that the subjects could correctly answer both questions for each target word on the comprehension test was .98 for the productive group, and .91 for the receptive group. While receptive recall must certainly be a factor in understanding a text, it is probably less of a factor than this result indicates. Since one aim of the design was to eliminate the many factors that may affect comprehension, the importance of vocabulary knowledge in the results is magnified. In a normal comprehension task, receptive recall is one of many factors that may affect the understanding of a text. While this study provides evidence that vocabulary knowledge may improve comprehension, in authentic learning the importance of vocabulary on comprehension is likely to depend on the type and number of factors involved.

The third question investigated by this research focused on the effect of pre-learning vocabulary on writing. The results indicate that L2 students that receive vocabulary instruction may be able to use some of the words in sentences. The subjects wrote an average of 35 percent of the words in correct sentences in the picture description test. This result is supported by L1 research that has shown that learners that receive vocabulary instruction may be able to successfully use recently taught words in their writing, and that their performance may depend on the method of learning (Duin and Graves, 1987; McDaniel and Pressley, 1984; Pressley, Levin, and Miller, 1981). The subjects that completed the productive learning from word pairs task used significantly

more of the target words in sentences (42%) than the receptive learning from word pairs group (29%). The difference between the two tasks indicates that productive tasks may have a greater effect on writing than receptive tasks. Since previous research has indicated that productive tasks may be effective (Duin and Graves, 1987), and that productive learning might be better suited to developing productive vocabulary knowledge than receptive learning (Experiment 3, Griffin and Harley, 1996; Stoddard, 1929; Waring, 1997b), it should not be surprising that productive learning from word pairs was superior on the picture description test. Writing is essentially a productive task that involves several different types of productive vocabulary knowledge. To write a sentence, learners must produce the forms of the words, and then use them with syntactic, semantic, and grammatical accuracy. Since receptive learning tends to focus learners on understanding language rather than producing it, it would be puzzling if receptive tasks were more effective than productive tasks in improving writing.

The results also suggest that relative gains in vocabulary knowledge may indicate how different tasks affect writing. The mean scores were higher on all of the productive vocabulary tests for the learners that completed the productive task than those that completed the receptive task although only orthography and meaning were significantly higher. If vocabulary knowledge can signal whether learners can use a word, perhaps productive recall may be a better indicator than receptive recall. This contrasts with McDaniel and Pressley's (1984) finding that recall of meaning may signal whether a learner can use a word correctly in a sentence. In the present study, the subjects could write sentences for 61 percent of the words for which they could recall the L2 form but only 54 percent of the words for which they could recall the L1 meaning. Although the relationship between use and receptive recall was higher in McDaniel and Pressley's L1 study, the similarity between test and task suggests that productive recall would be a bigger factor in writing while receptive recall would play a larger role in comprehension. It should be expected that the effects of meaning and form on writing would be lower for L2 learners than for L1 learners because L2 learners are likely to have lower ability using the language. It may not take much effort for a native speaker to use an unknown word when they know meaning and form but it may take

considerable effort for an L2 learner. Moreover, the relationship between productive recall and writing is likely to be less for learners of lower ability. Further research is needed to investigate the relationship between meaning and form and writing.

Taken as a whole, the results of the picture description and comprehension tests suggest that tests of vocabulary knowledge may provide some evidence of how tasks may affect communication. The results echoed the findings of previous studies showing that productive tasks may be better suited to developing productive knowledge, and receptive tasks may be more effective in developing receptive knowledge (Experiment 3, Griffin and Harley, 1996; Stoddard, 1929; Waring, 1997b). However, since very few studies have compared receptive and productive learning, additional research is also warranted.

In regard to instructional practice, the findings of this study may be informative for both teachers and learners. Those involved in vocabulary learning should be aware that the tasks that are used might have a powerful effect on what learners can and cannot do with a word. Since the majority of tasks used in vocabulary learning are receptive, they are well suited for improving comprehension but less appropriate for improving the use of taught words. Common teaching methods such as providing a definition or translation, and looking up words in the dictionary may be more conducive to increasing receptive knowledge and improving comprehension. However, if the aim of a learner or an instructional program is to improve speaking and writing, teachers and learners need to be aware that they may be more successful if they use productive tasks. If their goal is to improve overall language skills, the results indicate that a combination of receptive and productive tasks may prove to be most effective.

Teachers and learners should also be made aware that some decontextualized tasks may also be useful and effective. The trend in language teaching seems to have moved away from decontextualized to contextualized learning (Tinkham, 1989). However, the results of this study suggest that learning from word pairs may not only improve vocabulary learning but might also facilitate communicative skills such as

comprehension and writing. Since learning from word pairs is relatively fast, it could easily be incorporated into a language learning program.

5.8 Summary of main findings

The findings of Experiment 6 indicate:

1. L2 vocabulary learning may improve comprehension and writing with performance dependent on the method of instruction.
2. Receptive learning from word pairs was found to be a more effective method of increasing comprehension than productive learning from word pairs. In turn, this suggests that receptive learning may be better suited to improving comprehension than productive learning.
3. Productive learning from word pairs was found to be better suited to improving writing than receptive learning from word pairs. This suggests that productive tasks may be more effective if the aim is to use taught words.
4. Vocabulary knowledge may provide some indication of the efficacy of vocabulary learning tasks on writing.
5. Receptive recall may provide some evidence of whether vocabulary knowledge affects comprehension. However, since there are many other factors that affect comprehension the relationship between the two may vary between contexts.
6. Productive recall rather than receptive recall may signal whether learners can use a word correctly in a sentence.

CHAPTER 6

6.0 Introduction

This chapter will review the results of the six experiments which were presented in this thesis, and examine how the various vocabulary learning tasks affected vocabulary knowledge. This will be followed by a brief discussion on how the findings relate to testing and one recent theory of vocabulary acquisition - the Involvement Load Hypothesis.

6.1 Review of the findings

Experiment 1 investigated the effects of incidental learning from reading and direct learning of word pairs on word knowledge. The results showed that the learning from word pairs task contributed to surprisingly large gains in knowledge of syntax, and grammatical functions, as well as orthography, association, and meaning and form. The incidental learning task was also effective; however, gains were strongly influenced by the number of repetitions and the transparency of the contexts. Gains tended to increase as the number of repetitions increased; however, partial gains from an informative context may have been reduced or even eliminated if followed by a less informative or misleading context in the early stages of learning. The results indicated that it may not be possible to determine the number of repetitions needed to acquire a word because context has such a strong effect that the number may vary from word to word. This is what may

have occurred in the Saragi, Nation, and Meister (1978), and Horst, Cobb, and Meara (1998) studies. The results of the incidental task also indicated that gaining knowledge of orthography, association, syntax, and grammar may not be dependent on gaining knowledge of meaning since the learners were able to demonstrate knowledge of those aspects despite possessing relatively little knowledge of meaning.

One of the most interesting findings in the first experiment was the relative efficacy of the learning from word pairs task. Learners who did the word pairs task gained significantly more knowledge of all of the aspects than those who did the incidental learning from context task (1 repetition). Moreover, they also had significantly higher mean scores on the productive knowledge of syntax, and grammar tests than learners who met target words in seven contexts despite completing their task in a much shorter time. The learning from word pairs group was given four minutes, and the seven repetition group was given 28 minutes. However, since there were confounding variables in Experiment 1, the presence or absence of meaning and context, it is not clear how context affected vocabulary learning.

Overall, the results showed that both incidental learning from reading and learning from word pairs may be effective methods of gaining relatively full knowledge of a word. Learning from word pairs is likely to be the more effective task in the short term, and incidental learning may be more effective in the long term.

In Experiment 2 the only independent variable was context. Experiment 2 showed that both decontextualized (learning from word pairs) and contextualized learning (learning

from glossed sentences) contributed to large gains in all of the aspects of knowledge with no significant differences between the gains. This suggests that intermediate learners may be able to gain substantial L2 knowledge of syntax, grammar, and association if they gain knowledge of a word's L1 meaning. The overall findings also indicated that meeting unknown words in a single context may have little effect on gaining knowledge of syntax, and grammatical functions. However, when the words were classified as those with and without high frequency synonyms, the findings showed that the subjects gained significantly more knowledge of syntax from the contextualized task for the words without high frequency synonyms than in the decontextualized task. Since it was less likely that the subjects had L2 knowledge of syntax for the words without high frequency synonyms, they may provide a more accurate assessment of the effects of context. This result does provide evidence that context may increase learning; however, since there was no overall difference between the tasks, other factors such as learning burden may have an equally strong or stronger effect on acquisition.

The results of Experiment 2 gave support to Nation's (1990, 2001) suggestion that different words have different learning burdens. Experiment 2 indicated that synonyms may be easier to learn than non-synonyms. Both groups of subjects had superior gains for synonyms on the productive tests of syntax and association, and the learning from glossed sentences group also had higher scores on the receptive tests of grammar and orthography. The overlap in knowledge between unknown and known synonyms may reduce the learning burden and allow learners to use the known synonyms as models for the unknown synonyms. The overlap in syntax, grammar, and association may make the unknown words easier to learn than words such as non-synonyms that do not have as

high a degree of overlap. However, the gains in all of the aspects for non-synonyms suggest that intermediate L2 learners may be able to use their knowledge of L1 words as models for L2 counterparts. Again the degree of overlap, this time between L1 and L2 words, may determine the extent of the gains. Learning burden may provide an explanation for why this occurred.

Experiments 3, 4, 5, and 6 investigated the relative efficacy of receptive and productive learning. The results of Experiment 3 supported previous findings indicating that receptive learning may lead to higher gains in receptive knowledge while productive learning may lead to larger gains in productive knowledge (Griffin and Harley, 1996; Stoddard, 1929; Waring, 1997b). Learners who completed the productive task demonstrated significantly higher mean scores on four of the five productive tests, and the receptive test of orthography. They also scored higher on the productive measure of association but this result was not significant. The receptive group outperformed the productive group on all of the receptive tests except orthography; however, the only significant difference was on the test of meaning. The superiority of the productive task in promoting both receptive and productive orthographic knowledge suggests that productive learning may focus learners on form more than receptive learning.

The aim of Experiment 4 was to find how other receptive and productive tasks affect word knowledge. The results showed that both reading three glossed sentences and sentence production contributed to very large gains in all of the aspects of knowledge. However, the subjects that completed the receptive task had higher scores on all of the tests, and significantly higher scores on four of the five productive tests and the receptive

test of orthography. This contrasted with the findings in Experiment 3, as well as all of the earlier studies which had shown receptive learning to be more effective at increasing receptive knowledge and productive learning more effective at promoting productive knowledge (Griffin and Harley, 1996; Stoddard, 1929; Waring, 1997b). It was suggested that the ease with which receptive knowledge is gained, ceiling effects on the receptive measures, the time needed to complete the tasks, and the depth of processing involved in completing the tasks may account for the difference in results.

Experiment 5 involved the same tasks as Experiment 4 but in this study, time was considered to be a function of the task and therefore, was not controlled. The results again contrasted earlier findings (Experiment 3; Griffin and Harley, 1996; Stoddard, 1929; Waring, 1997b); however, this time the productive task was found to be significantly more effective on all 10 tests. The findings indicated that writing a sentence may be a much more effective task than reading three glossed sentences. However, it could also be argued that the writing task may have been superior because the learners may have spent more time writing than reading. Since the tasks in Experiment 5 were more typical of authentic vocabulary learning than those in Experiment 4, there may however, be a stronger argument in favor of sentence production. It is also important to note that the productive task was superior not only on the productive tests but also on the receptive tests. This suggests that if the aim of learning is to improve both receptive and productive knowledge, then productive learning may be most effective.

Experiment 6 investigated the effects of learning from receptive and productive word pairs on communication. A receptive test using a true/false format measured

comprehension, and a productive test using the picture description format measured writing. The results showed that the productive task contributed to higher scores on the writing test while the receptive task promoted superior scores on the comprehension test. This result expanded on earlier findings (Griffin and Harley, 1996; Stoddard, 1929; Waring, 1997b) as well as Experiment 3, and indicates that receptive learning may be better suited to improving receptive communication skills such as comprehension while productive learning may be superior in facilitating productive communication skills such as writing.

Although the primary aim of Experiment 6 was to investigate the effects of the tasks on communication, tests of vocabulary knowledge were also administered. The results of the vocabulary tests supported the findings of Experiment 3. The productive group scored higher on all of the productive tests, and significantly higher on the productive test of meaning and both tests of orthography. The receptive group scored higher on all of the receptive tests except for orthography but none of the differences were significant. The combined results of Experiments 3 and 6 indicated that productive learning from word pairs may be better suited to gaining both receptive and productive knowledge of orthography, and productive knowledge of meaning syntax, grammar, and association, while the receptive learning from word pairs task may be more effective in promoting receptive knowledge of meaning, association, grammar, and syntax. However, the lack of significant findings in favor of receptive learning in both experiments brings this into question, and suggests that further research is warranted. Overall, the results of Experiments 3 and 6 suggest that the word pairs tasks may enable intermediate students to gain a relatively comprehensive knowledge of target words in a very short time.

Taken as a whole, the results showed that all of the vocabulary learning tasks involved in the experiments may be effective. Each task contributed to gains in receptive and productive knowledge of orthography, meaning and form, association, syntax, and grammatical functions. However, the findings do show that different vocabulary learning tasks may result in different types of vocabulary knowledge. Experiments 3 and 6 indicated that receptive tasks may be better suited to developing receptive knowledge, and productive tasks may be more effective at promoting productive knowledge. While it was surprising that the contextualized tasks (incidental learning from reading and learning from glossed sentences) did not contribute to greater gains in knowledge of syntax, and grammar than the decontextualized task (learning from word pairs), this may have been due to the fact that subjects used in the experiments were at the intermediate level. Future research may show that contextualized tasks may be more effective at promoting gains in those aspects of knowledge with less advanced learners.

6.2 Testing

The primary difference between the experiments in this thesis and previous research was how word knowledge was measured. In the majority of studies, gains in knowledge have been measured by a single test of meaning (Read, 1987; Schmitt, 1994a). In all of the experiments in this thesis, 10 tests measuring knowledge of five aspects of vocabulary knowledge were administered. Orthography, meaning, syntax, association, and grammatical functions were each measured with two tests at different sensitivities. In Experiment 6, two additional tests (a true/false comprehension test, and a picture

description test) were administered to determine the communicative value of learning from receptive and productive word pairs. Taken as a whole, the results of the different tests provide some useful information on testing vocabulary.

First, measuring an aspect of knowledge with a single test does not always provide accurate results. In Experiments 1, 2, 3, 4, and 6, the results of the receptive tests often contrasted with the productive tests for a single aspect of knowledge. The studies repeatedly showed that one test may yield significant results while the other may not. This demonstrates the importance of using multiple tests of different sensitivities supporting Nagy, Herman, and Anderson's (1985) claims. Tests sensitive to small gains in knowledge are valuable because they are most likely to show that learning has occurred. However, if a test is too sensitive, there may be a ceiling effect which may decrease the likelihood of finding a significant difference. While a less sensitive test may also result in a floor effect, a combination of two tests at different sensitivities is more likely to show the extent of learning, and may provide more accurate results. Multiple tests may also reward researchers with a greater opportunity of finding significant results.

Second, tests that measure different aspects of knowledge may often have contrasting results. Experiments 1, 2, 3, 4, and 6 showed that gains may differ from one aspect of knowledge to another, and that while there may be little difference between the results for one aspect, another may yield significant findings. This suggests that when task effectiveness is being investigated, researchers should consider the effects of each task on the different aspects, and either test multiple aspects or test the aspect that is most

likely to result in a significant difference. Since very few studies have tested multiple aspects of knowledge this is one area of vocabulary research that needs to be explored further.

Third, assessing a single aspect of knowledge with both receptive and productive tests may be necessary to accurately measure vocabulary learning. All of the experiments in this thesis indicated that results may differ between receptive and productive tests for a single aspect of knowledge. Therefore, administering one type of test and not the other may produce misleading results. Experiments 3 and 6 supported previous research which indicated that receptive learning may be superior in developing receptive knowledge while productive learning may be better suited to increasing productive knowledge (Griffin and Harley, 1996; Stoddard, 1929; Waring, 1997b). Since there have been very few studies that have investigated the effects of different learning tasks on receptive and productive knowledge, further research is necessary. However, the results of this thesis strongly suggest that both receptive and productive tests should be administered if tasks involve both receptive and productive learning. Testing receptive knowledge may favor a receptive task while testing productive knowledge might favor a productive task. Moreover, measuring reception and production is more likely to show the extent of learning, particularly when comparing tasks that involve both types of learning, and it may also provide more accurate results.

Fourth, Experiment 6 in this thesis provided some evidence that vocabulary tests may signal communicative ability. It was found that the productive learning from word pairs tasks contributed to significantly higher scores on the picture description test and on the

productive tests of meaning and orthography. The mean scores for the productive group were also higher on the other productive tests; however, the differences were not significant. The receptive learning from word pairs task contributed to significantly higher scores on the comprehension test, and higher scores on the receptive tests of meaning, syntax, association, and grammatical functions; however, none of the differences on the vocabulary tests were significant. Since communication involves all five of the aspects that were tested, it might have been expected that the results on communicative tests would correspond with the results on the vocabulary tests. The lack of significant findings on the receptive tests may have been due to the experimental design or the receptive test format. However, it should also not be surprising that vocabulary tests such as the word association, translation and multiple-choice tests used in the experiments might not measure communicative ability. Standard vocabulary tests demonstrate what is known about words whereas communicative tests show what learners can or cannot do with those words. However, since the results offered some evidence that vocabulary tests may signal communicative ability, further research is needed to clarify this issue.

It is important for both teachers and students to be aware that there may be a difference between vocabulary tests and communicative tests. In many language classrooms time spent learning vocabulary may be limited. Standard vocabulary tests are a fast and efficient method of determining the extent of the students' learning. However, it would be wrong for students and teachers to believe that scoring correctly on a multiple choice vocabulary test ensures that students can communicate with those words. Using a combination of vocabulary and communicative tests may be the best method of

measuring knowledge and ability.

Fifth, there may be little need to use two correct answers as the criterion for a correct response on true/false tests. It was found in Experiment 6 that the subjects who scored correctly on one true/false question were extremely likely to score correctly on a second question about the same item. The correlation between scoring correctly on one question and scoring correctly on both questions about that item was .933. The rationale behind having learners answer two questions correctly about one item was to reduce the possibility of subjects scoring correctly through guessing from 50 percent to 25 percent. The size and significance of the correlation ($p < .000$) indicates that this may not be necessary. However, since very few studies have used true/false questions to assess word knowledge further research is warranted.

6.3 The Involvement Load Hypothesis

An area that has not been examined in this thesis is how the findings relate to current theory on vocabulary learning. One recent theory that has dealt with the relative efficacy of vocabulary learning tasks is the Involvement Load Hypothesis. Laufer and Hulstijn's (2001) Involvement Load Hypothesis suggests that L2 vocabulary acquisition is conditional upon the degree of involvement in processing unknown words. They identify three determining factors in vocabulary retention – need, search, and evaluation – and rate the degree of prominence of each factor for different learning tasks. The Involvement Load Hypothesis predicts that the higher the total rating for a task (involvement load), the more likely it is to be effective for vocabulary learning.

The need component refers to the motivation involved in vocabulary learning. When learning is intrinsically motivated such as when learners voluntarily look up a word in a dictionary during a task, it is rated as '2'. However, when the task requires the students to learn the word, need is rated as '1'. Search refers to the attempt to find the meaning or form of a word. If the meaning and form are given in a task, search is rated as '0', but if the learners have to check a dictionary, ask a teacher or peers, or search the text then it is rated as '1'. Evaluation refers to whether the word has to be compared or combined with other words. If learners have to compare different meanings (ie. cloze tasks), evaluation is rated as '1', but if a task requires learners to combine other words with an unknown word such as in sentence production, it is rated as '2'.

The need, search, and evaluation ratings are then totaled, and the combined rating (involvement load) is used to ascertain the relative efficacy of the tasks. For example, in a sentence production task in which the L1 meaning and L2 form are given, the need is rated as 1, search is 0, and evaluation is 2, for a total score of 3. However, in a learning from glossed sentences task, need is rated as 1, search is 0, and evaluation is 0, for a total of 1. Therefore, the higher involvement load for the sentence production task would indicate that it would be the more effective task. Laufer and Hulstijn (2001), and Hulstijn and Laufer (2001) have consistently found that tasks with a higher involvement load achieved superior results to those with lower involvement loads.

The Involvement Load Hypothesis is valuable because it provides a useful starting point to analyse L2 vocabulary tasks. The accuracy of the hypothesis in assessing the relative

efficacy of tasks in prior (Laufer and Hulstijn, 2001) and new (Hulstijn and Laufer, 2001) studies indicates that the three constructs involved – need, search, and evaluation – may be critical to L2 vocabulary acquisition. The results of the experiments in this thesis do, however, indicate that the hypothesis may not be entirely accurate, and may need modification. It should be noted that only the tasks from Experiments 1 and 5 meet Laufer and Hulstijn's (2001) criteria for incidental learning tasks since the learners were aware in the other experiments that they would be tested. However, it can be argued that incidental gains may have been made on all of the tests except for meaning and form, and orthography since the subjects were unaware of the nature and range of tests. Thus, all of the experiments may relate to the Involvement Load hypothesis.

Experiment 1 showed that two incidental learning tasks (learning from three informative contexts, and learning from three less informative contexts) are not equally effective despite having identical involvement loads. Experiments 3 and 6 suggest that similarity between the type of learning and the type of test (receptive and productive) may have a significant effect on the results. In those studies, receptive learning was found to be more effective when tested receptively while productive learning was superior when tested productively. Since the tasks had identical involvement loads, it would appear that similarity between direction of learning and the type of test may be a fourth factor that may need to be added to the hypothesis.

Experiment 1 also showed that the word pairs task was more effective than the incidental task despite having a lower involvement load. The incidental task's involvement load was 2 (need = 1, search = 1, and evaluation = 0) while the learning from word pairs task's

was 1 (need = 1, search = 0, and evaluation = 0). However, the results of Experiment 2 lend support to the hypothesis. The learning from glossed sentences task and the learning from word pairs tasks had the same involvement loads (need = 1, search = 0, and evaluation = 0), and no significant differences were found between the tasks on all ten tests.

The results of Experiment 4 contrast with the hypothesis. The reading task (need = 1, search = 0, and evaluation = 0) was found to be more effective than the writing task (need = 1, search = 0, and evaluation = 2) despite having a lower involvement load. However, since Hulstijn and Laufer (2001) report that they consider time on task to be an inherent property of a task, and the time given to complete the reading task was longer than the subjects needed, Experiment 5 may be a better indicator of the true effects of the tasks. Experiment 5 lends support to the involvement load hypothesis. The writing task had a higher involvement load than the reading task, and was found to be superior on all ten tests.

In summary, the findings in this thesis suggest that the Involvement Load Hypothesis may need to be modified. While it is a valuable first step towards building a theory of vocabulary acquisition, the results indicated that there may be more than three determining factors in vocabulary learning. The results of Experiment 1 suggested that context may have a significant affect on whether learning occurs, and the results of Experiments 3 and 6 indicated that the similarity between the direction of learning and the method of testing (receptive and productive) may also be a determining factor.

6.4 Conclusion

The purpose of this thesis has been to investigate the effects of vocabulary learning tasks on word knowledge. The experimental design expanded upon earlier methodologies by measuring vocabulary gains in knowledge of five different aspects with both receptive and productive tests. The results indicated that testing multiple aspects of knowledge may provide a better understanding of vocabulary learning tasks, testing, and vocabulary acquisition. The findings provided insight into the effects of the following vocabulary learning tasks:

1. Incidental vocabulary learning through reading (Experiment 1)
2. Learning from word pairs (Experiments 1, 2, 3, and 6)
3. Learning from glossed sentences (Experiments 2, 4, and 5)
4. Learning from receptive and productive word pairs (Experiments 3 and 6)
5. Learning from sentence production (Experiments 4 and 5)

The findings also showed how the following learning variables may affect vocabulary acquisition:

1. Repetition (Experiment 1)
2. Context (Experiments 1 and 2)
3. Synonymy and learning burden (Experiment 2)
4. Knowledge of meaning and form (Experiments 1-6)

Taken as a whole, this thesis shows the value of testing multiple aspects of word knowledge, and indicates that further research in this area is warranted. Since this is one of the first studies that has attempted to measure multiple aspects, there is much that remains to be learned. However, this thesis demonstrates that it is feasible to test five aspects receptively and productively. Since teachers are unlikely to have the time to create, administer, and assess gains in knowledge of a word with multiple tests, they may have to rely on researchers to measure the different aspects of knowledge. Past research has neglected to do this, hopefully future studies will not.

REFERENCES

- Aitchison, J. (1994) Words in the Mind (Second edition) Blackwell, Oxford.
- Bahns, J. (1993) Lexical collocations: a contrastive view. ELT Journal 47, 1: 56-63.
- Baker, L. (1989) Developmental changes in readers' responses to unknown words. Journal of Reading Behavior 21, 3: 241-260.
- Bamford, J. (1993) Beyond grammar translation: teaching students to really read. In P. Wadden (ed.) A handbook for teaching English at Japanese Colleges and Universities. Oxford University Press, Oxford: 63-72.
- Beck, I.L., McKeown, M.G. and McCaslin, E.S. (1983) Vocabulary: all contexts are not created equal. Elementary School Journal 83, 3: 177-181.
- Beck, I.L., McKeown, M.G. and Omanson, R.C. (1987) The effects and uses of diverse vocabulary instructional techniques. In McKeown and Curtis (eds.) The Nature of Vocabulary Acquisition Lawrence Erlbaum Associates, Mahwah, N.J.: 147-163.
- Beck, I.L., Perfetti, C.A. and McKeown, M.G. (1982) The effects of long-term vocabulary instruction on lexical access and reading comprehension. Journal of Educational Psychology 74: 506-521.
- Channell, J. (1988) Psycholinguistic considerations in the study of L2 vocabulary acquisition. In Carter and McCarthy (eds.) Vocabulary and Language Teaching Longman, London: 83-96.
- Craik, F.I.M. and Lockhart, R.S. (1972) Levels of processing: a framework for memory research. Journal of Verbal Learning and Verbal Behavior 11: 671-684.

- Craik, F.I.M. and Tulving, E. (1975) Depth of processing and the retention of words in episodic memory. Journal of Experimental Psychology 104: 268-294.
- Crow, J.T. (1986) Receptive vocabulary acquisition for reading comprehension. Modern Language Journal 70, 2: 242-250.
- Crow, J.T. and Quigley, J.R. (1985) A semantic field approach to passive vocabulary acquisition for reading comprehension. TESOL Quarterly 19, 3: 497-513.
- Cunningham, J.W. and Moore, D.W. (1993) The contribution of understanding academic vocabulary to answering comprehension questions. Journal of Reading Behavior 25, 2: 171-180.
- Day, R.R., Omura, C., and Hiramatsu, M. (1991) Incidental EFL vocabulary learning and reading. Reading in a Foreign Language 7, 2: 541-551.
- Dempster, F.N. (1987) Effects of variable encoding and spaced presentation on vocabulary learning. Journal of Educational Psychology 79, 2: 162-170.
- Duin, A.H. and Graves, M.F. (1986) Effects of vocabulary instruction used as a prewriting technique. Journal of Research and Development in Education 20, 1: 7-13.
- Duin, A.H. and Graves, M.F. (1987) Intensive vocabulary instruction as a prewriting technique. Reading Research Quarterly 22, 3: 311-330.
- Dupuy, B. and Krashen, S.D. (1993) Incidental vocabulary acquisition in French as a foreign language. Applied Language Learning 4, 1&2: 55-63.
- Eysenck, M.W. (1978) Levels of processing: a critique. British Journal of Psychology 69: 157-169.

- Gipe, J.P. and Arnold, R.D. (1979) Teaching vocabulary through familiar associations and contexts. Journal of Reading Behavior 11, 3: 282-285.
- Graves, M.F. (1986) Vocabulary learning and instruction. Review of Research in Education 13: 49-89.
- Griffin, G.F. (1992) Aspects of the psychology of second language vocabulary list learning. Unpublished PhD thesis, Dept of Psychology, University of Warwick.
- Griffin, G.F. and Harley, T.A. (1996) List learning of second language vocabulary. Applied Psycholinguistics 17: 443-460.
- Herman, P.A., Anderson, R.C., Nagy, W.E., and Pearson, P.D. (1987) Incidental acquisition of word meaning from expositions with varied text features. Reading Research Quarterly 22, 3: 263-284.
- Higa, M. (1965) The psycholinguistic concept of 'difficulty' and the teaching of foreign language vocabulary. Language Learning 15, 3&4: 167-179.
- Horst, M., Cobb, T., and Meara, P. (1998) Beyond a Clockwork Orange: acquiring second language vocabulary through reading. Reading in a Foreign Language 11, 2: 207-223.
- Hu, M. and Nation, I.S.P. (2000) Vocabulary density and reading comprehension. Reading in a Foreign Language 13, 1: 403-430.
- Hulstijn, J.H. (2001) Intentional and incidental second-language vocabulary learning: a reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (ed.) Cognition and Second Language Instruction Cambridge University Press, Cambridge.

- Hulstijn, J. and Laufer, B. (2001) Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. Language Learning 51, 3: 539-558.
- Hulstijn, J., Hollander, M. and Greidanus, T. (1996) Incidental vocabulary learning by advanced foreign language students: the influence of marginal glosses, dictionary use, and reoccurrence of unknown words. Modern Language Journal 80, 3: 327-339.
- Hulstijn, J.H. (1992) Retention of inferred and given word meanings: experiments in incidental vocabulary learning. In P. Arnaud and H. Bejoint (eds.) Vocabulary and Applied Linguistics Macmillan, London: 113-125.
- Jenkins, J.R., Pany, D., and Schreck, J. (1978) Vocabulary and reading comprehension: Instructional effects. Technical Report No. 100. Urbana: University of Illinois, Center for the Study of Reading.
- Jenkins, J.R., Matlock, B. and Slocumb, T.A. (1989) Two approaches to vocabulary instruction: the teaching of individual word meanings and practice in deriving word meanings from context. Reading Research Quarterly 24, 2: 215-235.
- Jenkins, J.R., Stein, M.L. and Wysocki, K. (1984) Learning vocabulary through reading. American Educational Research Journal 21, 4: 767-787.
- Joe, A. (1998) What effects do text-based tasks promoting generation have on incidental vocabulary acquisition? Applied Linguistics 19, 3: 357-377.
- Johnson, P. (1982) Effects on reading comprehension of building background knowledge. TESOL Quarterly 16, 4: 503-516.

- Kameenui, E.J., Carnine, D.C. and Freschi, R. (1982) Effects of text construction and instructional procedures for teaching word meanings on comprehension and recall. Reading Research Quarterly 17, 3: 367-388.
- Kellerman, E. (1985) If at first you do succeed... In S.M. Gass and C.G. Madden (eds.) Input in Second Language Acquisition Newbury House, Rowley, MA.: 345-353.
- Knight, S.M. (1994) Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal abilities. Modern Language Journal 78, 3: 285-299.
- Krashen, S.D. (1989) We acquire vocabulary and spelling by reading: additional evidence for the input hypothesis. Modern Language Journal 73, 4: 440-462.
- Kucera, H. and Francis, W.N. (1967) A computational analysis of present-day American English Brown University Press, Providence, R.I.
- Lado, R. (1956) Patterns of difficulty in vocabulary. Language Learning 6, 1&2: 23-41.
- Landauer, T.K. and Bjork, R.A. (1978) Optimum rehearsal patterns and name learning. In M.M. Gruneberg, P.E. Morris and R.N. Sykes (eds.) Practical Aspects of Memory Academic Press, London: 625-632.
- Laufer, B. (1998) The development of passive and active vocabulary: same or different? Applied Linguistics 19, 2: 255-271.
- Laufer, B. (1997) What's in a word that makes it hard or easy? Intralexical factors affecting the difficulty of vocabulary acquisition. In N. Schmitt and M. McCarthy (eds.) Vocabulary: Description, Acquisition and Pedagogy Cambridge University Press, Cambridge: 140-155.

- Laufer, B. (1992) How much lexis is necessary for reading comprehension? In P. Arnaud and H. Bejoint (eds.) Vocabulary and Applied Linguistics Macmillan, London: 126-132.
- Laufer, B. (1990a) Words you know: how they affect the words you learn. In J. Fisiak (ed.) Further Insights into Contrastive Linguistics. John Benjamins, Amsterdam: 573-593.
- Laufer, B. (1990b) Why are some words more difficult than others? Some intralexical factors that affect the learning of words. IRAL 28: 293-307.
- Laufer, B. (1990c) Ease and difficulty in vocabulary learning: some teaching implications. Foreign Language Annals 23, 2: 147-155.
- Laufer, B. (1989) A factor of difficulty in vocabulary learning: deceptive transparency. AILA Review 6: 10-20.
- Laufer, B. and Hulstijn, J. (2001) Incidental vocabulary acquisition in a second language: the construct of task-induced involvement. Applied Linguistics 22, 1: 1-26.
- Laufer, B. and Nation, P. (1999) A vocabulary size test of controlled productive ability. Language Testing 16, 1: 36-55.
- Laufer, B. and Paribakht, T.S. (1998) The relationship between passive and active vocabularies: effects of language learning context. Language Learning 48, 3: 365-391.
- Laufer, B. and Shmueli, K. (1997) Memorizing new words: Does teaching have anything to do with it? RELC Journal 28, 1: 89-108.

- Laufer, B. and Nation, P. (1995) Vocabulary size and use: lexical richness in L2 written production. Applied Linguistics 16, 3: 307-322.
- Liu Na and Nation, I.S.P. (1985) Factors affecting guessing vocabulary in context. RELC Journal 16, 1: 33-42.
- Ludwig, J. (1984) Vocabulary acquisition as a function of word characteristics. Canadian Modern Language Review 40, 5: 552-562.
- McCarthy, M. (1990) Vocabulary Oxford University Press, Oxford.
- McDaniel, M.A. and Pressley, M. (1984) Putting the keyword method in context. Journal of Educational Psychology 76: 598-609.
- McDaniel, M.A. and Pressley, M. (1989) Keyword and context instruction of new vocabulary meanings: effects on text comprehension and memory. Journal of Educational Psychology 81, 2: 204-213.
- McKeown, M.G. (1985) The acquisition of word meaning from context by children of high and low ability. Reading Research Quarterly 20: 482-496.
- McKeown, M.G., Beck, I.L., Omanson, R.C. and Perfetti, C. (1983) The effects of long-term vocabulary instruction on reading comprehension: a replication. Journal of Reading Behavior 15, 1: 3-18.
- McKeown, M.G., Beck, I.L., Omanson, R.G. and Pople, M.T. (1985) Some effects of the nature and frequency of vocabulary instruction on the knowledge and use of words. Reading Research Quarterly 20, 5: 522-535.
- Meara, P. (1996) The vocabulary knowledge framework [On-line]. Available Internet: <http://www.swan.ac.uk/cals/calsres/vlibrary/pm96d.htm>

- Meara, P. (1990) A note on passive vocabulary. Second Language Research 6, 2: 150-154.
- Meara, P. and Buxton, B. (1987) An alternative to multiple choice vocabulary tests. Language Testing 4, 2: 142-151.
- Melka, F. (1997) Receptive vs. productive aspects of vocabulary. In Schmitt and McCarthy (eds.) Vocabulary: Description, Acquisition and Pedagogy Cambridge University Press, Cambridge: 84-102.
- Melka Teichroew, F.J. (1982) Receptive vs. productive vocabulary: a survey. Interlanguage Studies Bulletin (Utrecht) 6, 2: 5-33.
- Mezynski, K. (1983) Issues concerning the acquisition of knowledge: effects of vocabulary training on reading comprehension. Review of Educational Research 53, 2: 253-279.
- Morgan, B.Q. and Oberdeck, L.M. (1930) Active and passive vocabulary. In E.W. Bagster-Collins (ed.) Studies in Modern Language Teaching 16, New York: 213-221.
- Morton, J. (1979) Word recognition. In J. Morton and J. Marshall (eds.) Structures and Processes 2, New York: Cornell University Press: 106-156.
- Miller, G.A. (1999) On knowing a word. Annual Review of Psychology 50: 1-19.
- Nagy, W.E. and Herman, P.A. (1987) Breadth and depth of vocabulary knowledge: implications for acquisition and instruction. In M. McKeown and M. Curtis (eds.) The Nature of Vocabulary Acquisition Lawrence Erlbaum Associates, Mahwah, N.J. : 19-35.

- Nagy, W.E. and Herman, P.A. (1985) Incidental vs. instructional approaches to increasing reading vocabulary. Educational Perspectives 23: 16-21.
- Nagy, W.E., Anderson, R.C. and Herman, P.A. (1987) Learning word meanings from context during normal reading. American Educational Research Journal 24, 2: 237-270.
- Nagy, W.E., Herman, P. and Anderson, R.C. (1985) Learning words from context. Reading Research Quarterly 20: 233-253.
- Nation, I.S.P. (2001) Learning Vocabulary in another Language. Cambridge University Press, Cambridge.
- Nation, I.S.P. (1990) Teaching and Learning Vocabulary Newbury House, MA.
- Nation, I.S.P. (1982) Beginning to learn foreign vocabulary: a review of the research. RELJ Journal 13, 1: 14-36.
- Nation, I.S.P. (1980) Strategies for receptive vocabulary learning. Guidelines 3: 18-23.
- Nation, P. and Wang, K. (1999) Graded readers and vocabulary. Reading in a Foreign Language 12, 2: 355-380.
- Nist, S.L. and Olejnik, S. (1995) The role of context and dictionary definitions on varying levels of word knowledge. Reading Research Quarterly 30, 2: 172-193.
- Oxford, R. and Crookall, D. (1990) Vocabulary learning: a critical analysis of techniques. TESL Canada Journal 7, 2: 9-30.
- Pany, D. and Jenkins, J.R. (1978) Learning word meanings: a comparison of instructional procedures. Learning Disability Quarterly 1: 21-32.

- Pany, D., Jenkins, J.R. and Schreck, J. (1982) Vocabulary instruction: effects on word knowledge and reading comprehension. Learning Disability Quarterly 5: 202-215.
- Paul, P.V., Stallman, A.C. and O'Rourke, J.P. (1990) Using three test formats to assess good and poor readers' word knowledge. Technical Report No. 509. Urbana: University of Illinois, Center for the Study of Reading.
- Pickering, M. (1982) Context-free and context-dependent vocabulary learning: an experiment. System 10, 1: 79-83.
- Pitts, M., White, H. and Krashen, S. (1989) Acquiring second language vocabulary through reading: a replication of the Clockwork Orange study using second language acquirers. Reading in a Foreign Language 5, 2: 271-275.
- Pressley, M., Levin, J.R. and McDaniel, M.A. (1987) Remembering versus inferring what a word means: mnemonic and contextual approaches. In M. McKeown and M. Curtis (eds.) The Nature of Vocabulary Acquisition Lawrence Erlbaum Associates, Mahwah, N.J.: 107-127.
- Pressley, M., Levin, J.R. and Miller, G.E. (1981) How does the keyword affect vocabulary comprehension and usage? Research Reading Quarterly 16: 213-226.
- Prince, P. (1996) Second language vocabulary learning: the role of context versus translations as a function of proficiency. Modern Language Journal 80, 4: 478-493.
- Read, J. (1993) The development of a new measure of L2 vocabulary knowledge. Language Testing 10, 3: 355-371.

- Read, J. (1987) Towards a deeper assessment of vocabulary knowledge. Paper presented at the 8th World Congress of Applied linguistics, Sydney, Australia.
- Richards, J.C. (1976) The role of vocabulary teaching. TESOL Quarterly 10, 1: 77-89.
- Rott, S. (1999) The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition through reading. Studies in Second Language Acquisition 21, 1: 589-619.
- Rystrom, R. (1970) Toward defining comprehension: A first report. Journal of Reading Behavior, 2: 56-74.
- Sackett, P.R. (1991) On interpreting measures of change due to training or other interventions: A comment on Casio (1989, 1991). Journal of Applied Psychology, 76, 4: 590-591.
- Saragi, T., Nation, I.S.P. and Meister, G.F. (1978) Vocabulary learning and reading. System 6, 2: 72-78.
- Schmitt, N. (2000) Vocabulary in Language Teaching Cambridge University Press, Cambridge.
- Schmitt, N. (1999) The relationship between TOEFL vocabulary items and meaning, association, collocation and word class knowledge. Language Testing 16, 2: 189-216.
- Schmitt, N. (1998) Tracking the incidental acquisition of second language vocabulary: a longitudinal study. Language Learning 48, 2: 281-317.
- Schmitt, N. (1994a) A fresh approach to vocabulary: using a word knowledge framework. Paper presented at RELC, 1994.

- Schmitt, N. (1994b) Vocabulary testing: questions for test development with six examples of tests of vocabulary size and depth. Thai TESOL Bulletin 6, 2: 9-16.
- Schmitt, N. (1993) Word knowledge: what it takes to really know a word. Unpublished paper.
- Schmitt, N., Schmitt, D. and Clapham, C. (2001) Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. Language Testing 18, 1: 55-88.
- Schmitt, N. and McCarthy, M. (eds.) (1997) Vocabulary: Description, Acquisition and Pedagogy Cambridge University Press, Cambridge.
- Schmitt, N. and Meara, P. (1997) Researching vocabulary through a word knowledge framework: word associations and verbal suffixes. Studies in Second Language Acquisition 19: 17-36.
- Seibert, L.C. (1930) An experiment on the relative efficiency of studying French vocabulary in associated pairs versus studying French vocabulary in context. Journal of Educational Psychology 21: 297-314.
- Shu, H., Anderson, R.C. and Zhang, Z. (1995) Incidental learning of word meanings while reading: a Chinese and American cross-cultural study. Reading Research Quarterly 30, 1: 76-95.
- Stahl, S.A. (1990) Beyond the instrumentalist hypothesis: some relationships between word meanings and comprehension. Technical report no. 505. Urbana: University of Illinois, Center for the Study of Reading.
- Stahl, S.A. and Fairbanks, M.M. (1986) The effects of vocabulary instruction: a model-based meta-analysis. Review of Educational Research 56, 1: 72-110.

- Stahl, S.A., Hare, V.C., Sinatra, R. and Gregory, J.F. (1991) Defining the role of prior knowledge and vocabulary in reading comprehension: the retiring of number 41. Journal of Reading Behavior 23, 4: 487-508.
- Stahl, S.A. and Jacobson, M.G. (1986) Vocabulary difficulty, prior knowledge, and text comprehension. Journal of Reading Behavior 18, 4: 309-323.
- Stahl, S.A., Jacobson, M.G., Davis, C.E. and Davis, R.L. (1989) Prior knowledge and difficult vocabulary in the comprehension of unfamiliar text. Reading Research Quarterly 24: 27-43.
- Sternberg, R.J. (1987) Most vocabulary is learned from context. In M. McKeown and M. Curtis (eds.) The Nature of Vocabulary Acquisition Lawrence Erlbaum Associates, Mahwah, N.J.: 89-105.
- Sternberg, R.J. and Powell, J.S. (1983) Comprehending verbal comprehension. American Psychologist 38: 878-893.
- Stoddard, G.D. (1929) An experiment in verbal learning. Journal of Educational Psychology 20: 452-457.
- Stratton, E.P. and Nacke, P.L. (1974) The role of vocabulary knowledge in comprehension. In P. Nacke (ed.) Interaction: reading and practice for college-adult reading National Reading Conference, Clemson, S.C.: 185-192.
- Thorndike, E.L. (1908) Memory for paired associates. Psychological Review 15: 122-138.
- Tinkham, T. (1993) The effect of semantic clustering on the learning of second language vocabulary. System 21, 3: 371-380.

- Tinkham, T. (1989) Rote learning, attitudes, and abilities: a comparison of Japanese and American students. TESOL Quarterly 23, 4: 695-698.
- Tuinman, J.J. and Brady, M.E. (1974) How does vocabulary account for variance on reading comprehension tests? A preliminary instructional analysis. In P. Nacke (ed.) Interaction: reading and practice for college-adult reading National Reading Conference, Clemson, S.C.: 176-184.
- van Daalen-Kapteijns, M.M. and Elshout-Mohr, M. (1981) The acquisition of word meanings as a cognitive learning process. Journal of Verbal Learning and Verbal Behavior 20: 386-399.
- Waring, R. (1999) Tasks for assessing second language receptive and productive vocabulary Unpublished PhD thesis, University of Wales.
- Waring, R. (1997a) The negative effects of learning words in semantic sets: a replication. System 25, 2: 261-274.
- Waring, R. (1997b) A study of receptive and productive learning from word cards. Studies in Foreign Languages and Literature Notre Dame Seishin University, Okayama 21, 1: 94-114.
- Waring, R. (1997c) A comparison of the receptive and productive vocabulary sizes of some second language learners. Immaculata Notre Dame Seishin University, Okayama 1: 53-68.
- Webb, W.B. (1962) The effects of prolonged learning on learning. Journal of Verbal Learning and Verbal Behavior 1: 173-182.
- Wixson, K.K. (1986) Vocabulary instruction and children's comprehension of basal stories. Reading Research Quarterly 21, 3: 317-329.

Wodinsky, M. and Nation, P. (1988) Learning from graded readers. Reading in a Foreign Language 5, 1: 155-161.

Xue Guoyi and Nation, I.S.P. (1984) A university word list. Language Learning and Communication 3, 2: 215-229.

Yoshikawa, M. (1998) Vocabulary acquisition by advanced learners: factors that cause difficulty for advanced learners. Australian Review of Applied Linguistics Series S Number 15: 169-182.

APPENDIX A

Experiment 1 incidental learning contexts.

Repetition 1

He was not ill, and of course the beds in the ancon are for ill people.

He can read and faddam, and he thinks a lot.

And in London, hundreds of people were waiting at King's Cross Station for the masco from Harrogate.

He insisted on paconing a bright yellow sports jacket and diamond tie-pin while serving lunch.

I don't want to be rich, but I do want to come to Australia and sagod you and Mollie and my grandchildren.

Don't you denent you promised you would never leave me?

After paying the driver and getting out, he suddenly realized how dark and dirty the narrow dangies were.

He had cader and tea with Mrs. Walsh in the kitchen, and he felt really happy.

As she stood up to see if her bag was all right, she saw her hodet in the mirror - white, with big, round eyes.

Archie stayed for the rest of the afternoon, and for supper that tasper.

(Experiment 1 incidental learning contexts continued)

Repetition 2

Perhaps I had a Christmas with my mother once, but I do not denent it.

So John had to walk back through the dangies in the middle of the night, all wet from the sea, while we laughed at him from the car.

The first time Mary and the children laughed together, John felt a big smile come onto his hodet.

Time passed slowly in the ancon, where the patients played cards, and slept, and told each other stories.

She was paconing a white coat with a lot of pens in the top pocket.

I sagoded him every day, and talked to him.

I've ordered some champagne and some cader for us.

There are no long taspers in our New Zealand days; the sun goes down and half an hour later it's night.

She asked a lot of people a lot of questions and always faddamed the answers in her notebook.

Agatha still only eighteen years old went to London on the masco.

(Experiment 1 incidental learning contexts continued)

Repetition 3

Children played in the narrow dangles.

On her last day, Mary and John had cader in the factory restaurant together.

She remembered seeing a beautiful young girl in a hotel in Cairo when she was sagoding Egypt with Clara.

We liked to talk to him, and we are all very sorry because he is dead. A lot of people are going to denent him for a long time.

Their friends were paconing white clothes, with Greenworld written on them.

I could not see his hodet or his body.

Bob Tappin and Bob Leeming played music in the hotel each tasper, and both of them watched the quiet woman in the corner of the room and began to think.

I am a doctor at the London Ancon.

Your brother faddamed his name on this paper.

The masco began to go more slowly. It gave a long whistle. They were coming to a town.

(Experiment 1 incidental learning contexts continued)

Repetition 4

Yes, pacon your white dress.

During the summer of 1915, Agatha was ill and could not do any nursing work at the ancon for three or four weeks.

Don't leave your seat on the masco except to go and wash your hands, and when you do that, make sure you lock the door.

Then, in the autumn of 1928, Agatha decided to sagod the West Indies.

Where did the idea for the character come from? Where? 'I can never denent,' Agatha always said.

He's asked me to faddam about the environment for our newspaper.

On Monday tasper it rained again.

The grandfather put up his umbrella, and held it over her as they walked to a restaurant to have cader.

How big her eyes look in that white hodet, he thought.

They put all the bags in the car and drove away. Karen stood alone in the middle of the dangie.

(Experiment 1 incidental learning contexts continued)

Repetition 5

After cader they went to the English Garden.

At about eleven o'clock that tasper, Agatha went out and drove away in her car. She did not return home that night.

She came back to the town sometimes, and twice he had sagoded her in Scotland.

Has he no family, no old mother, who will always denent his face?

'Do you know, there's a young man just like your friend Potts down in the dangie,' she said from the window.

'She's ill, that's all,' I said quickly. 'We're going to the ancon.'

He was a small, uninteresting looking man in grey clothes, with grey hair and a thin hodet.

Paul's old school tie, which he was paconing, looked very similar to the Bollinger Club tie.

She's only thirteen years old, but she can read and faddam in four languages already.

There is an excellent masco that leaves London at about ten.

(Experiment 1 incidental learning contexts continued)

Repetition 6

My wife was a famous sailor . . . Rachel Horsley ... Perhaps you
denent her.

All the people in the dangie were very angry about the noise.

John's masco was late, and he caught a taxi from the station.

My lady insisted on my paconing a proper uniform from the start.

Her hodet was very white, and he had no idea if her heart was beating.

He could read about things, and talk to his visitors, but he could not go out of the ancon
by himself. He thought and played like a child.

The doctor began to faddam something on a piece of paper.

Nobody knows what Agatha was thinking, late that dark winter tasper.

By two o'clock they had finished cader, and Margot's second-best Rolls Royce was
waiting outside the hotel.

'She's invited me to sagod her in London, you know.'

(Experiment 1 incidental learning contexts continued)

Repetition 7

She will arrive at six the same tasper, so you can have a nice quiet day to rest and practice your German.

'You're going by masco on Saturday,' Mrs Doolin said.

She saw, heard, and denented everything - names, faces, the times of trains and buses, the colour of a shirt, the sound of a door shutting.

They looked for it at the ancon and at the police station.

I have to faddam a full page article every week on the environment.

'Ah, Uncle,' I said. 'Let's go in and have some cader.'

She was older than Wilson – about thirty five, perhaps - with short brown hair, and a pretty, friendly hodet.

There, quite by chance, the first thing she saw was a lovely girl in the mirror, paconing a beautiful black and gold hat.

They were like bright birds that had come to sagod the Sheridans' garden for this one afternoon, on their way to - where?

She walked along the dangie, asking herself those questions.

(Experiment 1 incidental learning contexts continued)

Repetition 8

After that evening he began to send me presents, and he asked me to sagod him at Hampton Court Palace.

I've always wanted to ride on that famous masco.

Everyone pacons special clothing.

Think how many times the sun has gone down in the taspers.

He ran down the room, out of the door, into the dangie.

Just before cader a young man arrived at the Castle, with a little pile of clothes he had found on the beach.

She saw how tired and sad his hodet was.

One day in 1884, I saw a picture in the window of a shop near the ancon.

I'd like to try and faddam a detective story myself,' said Agatha.

He denents his mother, and he has a picture of her. She was beautiful and kind, he says. But he never sees her now.

(Experiment 1 incidental learning contexts continued)

Repetition 9

They were all paconing white gloves and their hands were inside a big glass box.

At the corner of the dangie she met Laurie.

I did not talk to him very much at the ancon. I looked at his head and arms and legs and body very carefully.

They danced together many times that tasper, and Archie told her his plans.

Then a cat came to sagod me - a beautiful cat - and then a dog.

Then Simon stood up to say something. 'Mr Duncan,' he said. 'I'll always denent this night.

Everyone on the masco and the ship looked at him, and laughed at him.

Paul found the Minister sitting alone in the garden after cader, smoking a large cigar, his big red hands folded over his large stomach, a soft hat over his eyes.

'Why don't you faddam another book?' said Archie.

His hodet was red and he looked at his shoes.

(Experiment 1 incidental learning contexts continued)

Repetition 10

I want to read what she faddamed to her dear husband.

As soon as he could walk, he left the ancon and started looking for a ship to take him back to England.

After cader we sat and talked for a while.

His hodet was shining, excited, happy.

Her brother Edward was always with her when the King came to sagod.

Perhaps this happened. Or perhaps she's ill and can't denent who she is.

During the tasper, she talked to some people who were just back from Baghdad, in Iraq.

Big white clouds over the white dangies - and sunshine everywhere.

The masco left the station and rushed into the dark.

I don't know anything about art, and I haven't met any grand people, and I don't pacon expensive clothes, but up to now, none of that has worried me.

APPENDIX B

Experiment 1 word pairs.

| | |
|--------|----------|
| ancon | hospital |
| faddam | write |
| masco | train |
| pacon | wear |
| sagod | visit |
| denent | remember |
| dangy | street |
| cader | lunch |
| hodet | face |
| tasper | evening |

APPENDIX C

Experiment 1 vocabulary knowledge tests

In the experiments each test appeared on a separate page.

Test 1.

You will hear each word that you learned twice. Spell the words correctly.

1. _____
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____
9. _____
10. _____

Test 2.

Which spelling is correct?

- | | | | | |
|-----|-----------|-----------|-----------|-----------|
| 1. | a) denant | b) danant | c) danent | d) denent |
| 2. | a) tospar | b) tesper | c) tasper | d) tespar |
| 3. | a) hodet | b) holat | c) halet | d) hedet |
| 4. | a) codar | b) cader | c) cadar | d) ceder |
| 5. | a) dengie | b) dengy | c) dongie | d) dangy |
| 6. | a) fadam | b) faddem | c) fadem | d) faddam |
| 7. | a) ancon | b) ankon | c) anken | d) ancan |
| 8. | a) pacan | b) pacen | c) pacon | d) pacun |
| 9. | a) sagad | b) saggud | c) saggod | d) sagod |
| 10. | a) masco | b) masko | c) musko | d) musco |

(Experiment 1 vocabulary knowledge tests continued)

Test 3

Write the Japanese translation of each word.

For example, person
exit

人
出口

1. faddam _____
2. dangy _____
3. tasper _____
4. pacon _____
5. sagod _____
6. denent _____
7. ancon _____
8. cader _____
9. hodet _____
10. masco _____

Test 4

Write each word in a sentence.

For example, man
eat
chair

I saw a tall man talking to a woman.
I want to eat fried chicken.
It is a chair.

1. hodet _____
2. pacon _____
3. cader _____
4. masco _____
5. ancon _____
6. sagod _____
7. tasper _____
8. dangy _____
9. faddam _____
10. denent _____

(Experiment 1 vocabulary knowledge tests continued)

Test 5

Write a word which is often together with the word in a sentence.

次の単語は、どんなことばと一緒に使われますか？

For example, play friends
play baseball

1. sagod _____
2. tasper _____
3. cader _____
4. masco _____
5. pacon _____
6. denent _____
7. faddam _____
8. dangy _____
9. hodet _____
10. ancon _____

Test 6

Write a word which has a relationship with the word (similar meaning, opposite meaning, related meaning).

次の単語に関係のある言葉を書きなさい。(同意語, 反意語, 関連語など)

For example, man woman
hot temperature
winter snow

1. masco _____
2. faddam _____
3. denent _____
4. tasper _____
5. sagod _____
6. ancon _____
7. dangy _____
8. cader _____
9. pacon _____
10. hodet _____

(Experiment 1 vocabulary knowledge tests continued)

Test 7

Which sentence is correct?

1. a) It is an ancon.
b) He anconed.
c) It is very ancon.
2. a) I denented him.
b) It is a denent.
c) It is very denent.
3. a) It is a masco.
b) It mascoed.
c) It is very masco.
4. a) It is very sagod..
b) It is a sagod.
c) I sagoded her.
5. a) It cadered.
b) It is very cader.
c) It is my cader.
6. a) It is quiet in the tasper.
b) They taspered quietly.
c) It is very tasper.
7. a) It is a pacon.
b) It is very pacon.
c) I paconed it.
8. a) It is very faddam.
b) I faddamed him.
c) It is a faddam.
9. a) They are very hodet.
b) I saw their hodets.
c) They are going to hodet.
10. a) It dangied.
b) It is really dangy.
c) It is a dangy.

(Experiment 1 vocabulary knowledge tests continued)

Test 8

Which words are often together in a sentence?

文章にすると、組み合わせて使われる可能性の高い単語を選びなさい

For example, play friends
play baseball

- | | | | | |
|-----------|-------------|-----------|--------------|------------------|
| 1. faddam | a) letter | b) dirty | c) bag | d) pay |
| 2. hodet | a) town | b) mirror | c) leave | d) serve |
| 3. dangy | a) round | b) talk | c) happy | d) narrow |
| 4. ancon | a) tea | b) ill | c) afternoon | d) give |
| 5. denent | a) promise | b) sun | c) driver | d) yellow |
| 6. sagod | a) feel | b) stand | c) dark | d) grandchildren |
| 7. masco | a) show | b) think | c) bag | d) station |
| 8. cader | a) kitchen | b) white | c) round | d) rich |
| 9. tasper | a) people | b) eyes | c) supper | d) talk |
| 10. pacon | a) suddenly | b) bed | c) jacket | d) wait |

Test 9

Which word is associated with each word (similar meaning, opposite meaning, related meaning)?

関係のある単語を選びなさい。(同意語、反意語、関連語など)

For example, man woman
hot temperature
winter snow

- | | | | | |
|------------|--------------|-----------------|--------------|-------------------|
| 1. ancon | a) teacher | b) businessman | c) policeman | d) nurse |
| 2. faddam | a) read | b) play | c) sleep | d) touch |
| 3. masco | a) work | b) conversation | c) shopping | d) transportation |
| 4. pacon | a) education | b) winter | c) fashion | d) art |
| 5. sagod | a) listen | b) travel | c) sit | d) think |
| 6. denent | a) buy | b) welcome | c) forget | d) clean |
| 7. dangy | a) glass | b) sky | c) town | d) road |
| 8. cader | a) food | b) flowers | c) people | d) paper |
| 9. hodet | a) telephone | b) bag | c) body | d) desk |
| 10. tasper | a) family | b) morning | c) piano | d) money |

(Experiment 1 vocabulary knowledge tests continued)

Test 10

Choose the correct meaning for each word.

- | | | | | |
|------------|-----------|-------------|-------------|----------------|
| 1. ancon | a) house | b) car | c) city | d) hospital |
| 2. faddam | a) write | b) drink | c) sleep | d) walk |
| 3. masco | a) train | b) table | c) mountain | d) window |
| 4. pacon | a) speak | b) cry | c) laugh | d) wear |
| 5. sagod | a) buy | b) listen | c) visit | d) sit |
| 6. denent | a) hold | b) remember | c) jump | d) shout |
| 7. dangy | a) store | b) street | c) door | d) office |
| 8. cader | a) school | b) lunch | c) movie | d) businessman |
| 9. hodet | a) bottle | b) book | c) clock | d) face |
| 10. tasper | a) month | b) chair | c) evening | d) bath |

APPENDIX D

Experiment 2 word pairs.

Learn the 20 words (下記の英単語(20語)を覚えなさい。)

Version 1

| | |
|---------------|--------|
| ancon | 嘆く |
| gishom | 短剣 |
| dangy | マリネにする |
| nuggy | 巨石 |
| faddam | 質に入れる |
| mesut | 住居 |
| hodet | なめる |
| dapew | 小道 |
| denent | 礁 |
| ictay | 容貌 |
| masco | 槍 |
| tagon | 機関車 |
| pacon | 世捨て人 |
| nasin | むせび泣く |
| cader | ごみ廃棄場 |
| hattaw | うとうとする |
| tasper | 修道院 |
| toncop | 熱望する |
| sagod | 埠頭 |
| copac | ひどく嫌い |

Version 2

| | |
|---------------|--------|
| copac | 嘆く |
| ancon | 短剣 |
| gishom | マリネにする |
| dangy | 巨石 |
| nuggy | 質に入れる |
| faddam | 住居 |
| mesut | なめる |
| hodet | 小道 |
| dapew | 礁 |
| denent | 容貌 |
| ictay | 槍 |
| masco | 機関車 |
| tagon | 世捨て人 |
| pacon | むせび泣く |
| nasin | ごみ廃棄場 |
| cader | うとうとする |
| hattaw | 修道院 |
| tasper | 熱望する |
| toncop | 埠頭 |
| sagod | ひどく嫌い |

APPENDIX E

Experiment 2 glossed sentences.

Version 2

Learn the 20 words. (下記の英単語(20語)を覚えなさい。)

copac 嘆く They continue to copac for years after the death of their friend.

ancon 短剣 She picked up his ancon when he wasn't looking and cut him.

gishom マリネにする He picked up his fork and looked at the small fish
gishomed in olive oil and garlic.

dangy 巨石 The dangy was as large as a small house.

nuggy 質に入れる He nuggied his watch to buy some new clothes.

faddam 住居 He had decorated his new faddam with the finest furniture.

mesut なめる The dog jumped up and mesuted his face.

hodet 小道 He walked them out to the big car at the end of the hodet.

dapew 礁 The small boat went south around the dapew.

denent 容貌 He was always careful to wash his denent.

ictay 槍 He was killed with the long hunting ictay.

masco 機関車 The driver got off the masco.

tagon 世捨て人 It was true that he was a tagon and never came to the village.

pacon むせび泣く She stopped screaming and began to pacon.

nasin ごみ廃棄場 Most of this garbage goes to nasin—very large holes in the ground
far from where most people live.

cader うとうとする Closing her eyes she cadered.

hattaw 修道院 She doesn't want to stay inside a hattaw for the rest of her life.

tasper 熱望する I used to tasper expensive clothes, and jewelry as well.

toncop 埠頭 My brother and I were at the end of the toncop, fishing.

sagod ひどく嫌い We do **sagod** dust and dirt, and dirt in the bathtub.

APPENDIX F

Experiment 2 vocabulary knowledge tests.

Version 2

Test 1.

You will hear each word that you learned twice. Spell the words correctly.

テープから流れる単語を聞き取りなさい。(同じ単語が 2 回流れます。)

1. _____
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____
9. _____
10. _____
11. _____
12. _____
13. _____

14. _____
15. _____
16. _____
17. _____
18. _____
19. _____
20. _____

(Experiment 2 vocabulary knowledge tests continued)

Test 2

Which spelling is correct? どのつづりが正しいですか？

- | | | | | |
|-----|-----------|-----------|-----------|-----------|
| 1. | a) denant | b) danant | c) danent | d) denent |
| 2. | a) tospar | b) tesper | c) tasper | d) tespar |
| 3. | a) hodet | b) holat | c) halet | d) hedet |
| 4. | a) codar | b) cader | c) cadar | d) ceder |
| 5. | a) dengie | b) dengy | c) dungie | d) dangy |
| 6. | a) fadam | b) faddem | c) fadem | d) faddam |
| 7. | a) ancon | b) ankon | c) anken | d) ancan |
| 8. | a) pacan | b) pacen | c) pacon | d) pacun |
| 9. | a) sagad | b) saggud | c) saggod | d) sagod |
| 10. | a) masco | b) masko | c) musko | d) musco |
| 11. | a) nasin | b) nesen | c) nasen | d) nesin |
| 12. | a) copak | b) copack | c) copauk | d) copac |

- | | | | | |
|-----|-----------|-----------|-----------|-----------|
| 13. | a) ickty | b) ickay | c) ictay | d) ictey |
| 14. | a) nogey | b) noggy | c) nuggy | d) nugey |
| 15. | a) hattaw | b) hathaw | c) hathow | d) hattow |
| 16. | a) tagen | b) togen | c) togon | d) tagon |
| 17. | a) gishom | b) gissam | c) gisham | d) gashom |
| 18. | a) tancap | b) tancop | c) toncap | d) toncop |
| 19. | a) masut | b) masot | c) mesut | d) mesot |
| 20. | a) dopew | b) dapew | c) depew | d) dipew |

(Experiment 2 vocabulary knowledge tests continued)

Test 3

Write the words you learned in Part 1 beside the Japanese translations. Do not write words that were not in Part 1. 覚えた単語を書きなさい。

1. 嘆く _____
2. 短剣 _____
3. マリネにする _____

- | | | |
|-----|--------|-------|
| 4. | 巨石 | _____ |
| 5. | 質に入れる | _____ |
| 6. | 住居 | _____ |
| 7. | なめる | _____ |
| 8. | 小道 | _____ |
| 9. | 礁 | _____ |
| 10. | 容貌 | _____ |
| 11. | 槍 | _____ |
| 12. | 機関車 | _____ |
| 13. | 世捨て人 | _____ |
| 14. | むせび泣く | _____ |
| 15. | ごみ廃棄場 | _____ |
| 16. | うとうとする | _____ |
| 17. | 修道院 | _____ |
| 18. | 熱望する | _____ |
| 19. | 埠頭 | _____ |
| 20. | ひどく嫌い | _____ |

(Experiment 2 vocabulary knowledge tests continued)

Test 4

Write each word in a sentence.

与えられた単語を使って英文を作りなさい。

For example, man I saw a tall man talking to a woman.

 eat I want to eat fried chicken.

 chair It is a chair.

1. hodet _____
2. hattaw _____
3. pacon _____
4. ictay _____
5. cader _____
6. copac _____
7. masco _____
8. gishom _____
9. sagod _____
10. nuggy _____
11. tasper _____
12. dapew _____
13. dangy _____
14. mesut _____
15. faddam _____
16. nasin _____
17. denent _____

18. tagon _____

19. ancon _____

20. toncop _____

(Experiment 2 vocabulary knowledge tests continued)

Test 5

Write a word which is often together with the word in a sentence.

次の単語は、どんなことばと一緒に使われますか？

For example, play friends

play baseball

1. sagod _____

2. toncop _____

3. tasper _____

4. hattaw _____

5. cader _____

6. nasin _____

7. pacon _____

8. tagon _____

9. masco _____

10. ictay _____

11. denent _____
12. dapew _____
13. hodet _____
14. mesut _____
15. faddam _____
16. nuggy _____
17. dangy _____
18. gishom _____
19. ancon _____
20. copac _____

(Experiment 2 vocabulary knowledge tests continued)

Test 6

Write a word which has a relationship with the word (similar meaning, opposite meaning, related meaning).

次の単語に関係のある言葉を書きなさい。(同意語, 反意語, 関連語など)

For example, man woman
 hot temperature
 winter snow

1. copac _____

2. masco _____
3. mesut _____
4. faddam _____
5. nuggy _____
6. denent _____
7. dapew _____
8. tasper _____
9. ictay _____
10. sagod _____
11. gishom _____
12. ancon _____
13. hattaw _____
14. dangy _____
15. toncop _____
16. cader _____
17. nasin _____
18. pacon _____
19. tagon _____
20. hodet _____

(Experiment 2 vocabulary knowledge tests continued)

Test 7

Which sentence is correct? 正しい文章を選びなさい。

- | | | | |
|-----|-----------------------|------------------------|-----------------------|
| 1. | a) It is an ancon. | b) He anconed. | c) It is very ancon. |
| 2. | a) He is copacing. | b) He is a copac. | c) He is very copac. |
| 3. | a) I denented her. | b) It is her denent. | c) It is very denent. |
| 4. | a) She is a nuggy. | b) She nuggied it. | c) She is very nuggy. |
| 5. | a) It is a masco. | b) It mascoed. | c) It is very masco. |
| 6. | a) He dapewed it. | b) It is the dapew. | c) It is very dapew. |
| 7. | a) It is very sagod. | b) She is a sagod. | c) I sagoded her. |
| 8. | a) He is very mesut. | b) He mesuted it. | c) He is a mesut. |
| 9. | a) He cadered. | b) He is very cader. | c) He is my cader. |
| 10. | a) It toncoped. | b) It is a toncop. | c) It is very toncop. |
| 11. | a) He is a tasper. | b) He taspered it. | c) He is very tasper. |
| 12. | a) It is a hattaw. | b) It will hattaw. | c) It is very hattaw. |
| 13. | a) She is a pacon. | b) She is very pacon. | c) She paconed. |
| 14. | a) He has a gishom. | b) He gishomed it. | c) He is very gishom. |
| 15. | a) It is very faddam. | b) I faddamed it. | c) It is my faddam. |
| 16. | a) It is an ictay. | b) It is very ictay. | c) It ictayed. |
| 17. | a) It is very hodet. | b) It is a hodet. | c) It hodeted. |
| 18. | a) She is the tagon. | b) She tagoned. | c) She is very tagon. |
| 19. | a) It dangied. | b) It is really dangy. | c) It is a dangy. |
| 20. | a) It is very nasin. | b) It nasined. | c) It is a nasin. |

(Experiment 2 vocabulary knowledge tests continued)

Test 8

Which words are often together in a sentence?

文章にするとき、組み合わせて使われる可能性の高い単語を選びなさい

For example, play friends

play baseball

- | | | | | |
|-----------|---------------|---------------|-------------|--------------|
| 1. mesut | a) ice-cream | b) steak | c) potatoes | d) cake |
| 2. hodet | a) drive | b) sit | c) take | d) know |
| 3. nuggy | a) beautiful | b) fashion | c) model | d) jewelry |
| 4. dangy | a) fall | b) wash | c) walk | d) catch |
| 5. gishom | a) dessert | b) restaurant | c) meat | d) ice-cream |
| 6. ancon | a) wood | b) metal | c) plastic | d) stone |
| 7. copac | a) traveled | b) died | c) fished | d) bought |
| 8. denent | a) new | b) pretty | c) hungry | d) plastic |
| 9. hattaw | a) restaurant | b) church | c) office | d) house |
| 10. sagod | a) water | b) snakes | c) cars | d) flowers |
| 11. nasin | a) delicious | b) dirty | c) soft | d) pretty |
| 12. masco | a) leaves | b) falls | c) washes | d) cuts |
| 13. tagon | a) buy | b) walk | c) eat | d) become |
| 14. cader | a) sad | b) excited | c) tired | d) hungry |

- | | | | | |
|------------|---------------|-----------|------------|-------------|
| 15. ictay | a) cook | b) play | c) jump | d) throw |
| 16. tasper | a) ice-cream | b) school | c) dinner | d) money |
| 17. toncop | a) city | b) church | c) ocean | d) building |
| 18. pacon | a) water | b) tears | c) ocean | d) oil |
| 19. dapew | a) fish | b) beef | c) salad | d) dinner |
| 20. faddam | a) vegetables | b) dinner | c) flowers | d) garden |

(Experiment 2 vocabulary knowledge tests continued)

Test 9

Which word is associated with each word (similar meaning, opposite meaning, related meaning)?

関係のある単語を選びなさい。(同意語, 反意語, 関連語など)

For example, man woman

hot temperature

winter snow

- | | | | | |
|----------|------------|------------|-----------|-----------|
| 1. ancon | a) plate | b) spoon | c) knife | d) fork |
| 2. mesut | a) nose | b) ears | c) mouth | d) eyes |
| 3. masco | a) travel | b) work | c) study | d) like |
| 4. ictay | a) country | b) America | c) Europe | d) Africa |
| 5. sagod | a) love | b) know | c) want | d) need |
| 6. tagon | a) friends | b) date | c) party | d) alone |

| | | | | |
|------------|-------------|---------------|------------|--------------|
| 7. dangy | a) stone | b) plant | c) tree | d) person |
| 8. nasin | a) relax | b) work | c) study | d) clean |
| 9. hodet | a) park | b) highway | c) garden | d) building |
| 10. gishom | a) work | b) wash | c) clean | d) cook |
| 11. faddam | a) office | b) restaurant | c) store | d) house |
| 12. hattaw | a) religion | b) love | c) peace | d) happiness |
| 13. pacon | a) bored | b) excited | c) angry | d) sad |
| 14. dapew | a) food | b) wine | c) oil | d) water |
| 15. denent | a) shopping | b) fashion | c) body | d) clothes |
| 16. nuggy | a) buy | b) sell | c) want | d) like |
| 17. cader | a) sleep | b) eat | c) go | d) buy |
| 18. copac | a) party | b) funeral | c) wedding | d) baby |
| 19. tasper | a) like | b) want | c) hate | d) have |
| 20. toncop | a) car | b) ship | c) train | d) bus |

(Experiment 2 vocabulary knowledge tests continued)

Test 10

Write the Japanese translation of each word using words that you learned today.

次の単語を日本語に直しなさい。

For example, person 人 _____

exit 出口 _____

1. gishom _____
2. faddam _____
3. hattaw _____
4. dangy _____
5. mesut _____
6. tasper _____
7. ictay _____
8. pacon _____
9. nasin _____
10. sagod _____
11. dapew _____
12. denent _____
13. toncop _____
14. ancon _____
15. nuggy _____
16. cader _____
17. tagon _____
18. hodet _____
19. copac _____
20. masco _____

APPENDIX G

Experiment 3 receptive word pairs.

| | |
|--------|--------|
| masco | 機関車 |
| denent | 容貌 |
| hodet | 小道 |
| faddam | 住居 |
| dangy | 巨石 |
| tasper | 熱望する |
| cader | うとうとする |
| pacon | むせび泣く |
| sagod | ひどく嫌い |
| ancon | 短剣 |

APPENDIX H

Experiment 3 productive word pairs.

| | |
|--------|--------|
| 機関車 | masco |
| 容貌 | denent |
| 小道 | hodet |
| 住居 | faddam |
| 巨石 | dangy |
| 熱望する | tasper |
| うとうとする | cader |
| むせび泣く | pacon |
| ひどく嫌い | sagod |
| 短剣 | ancon |

APPENDIX I

Experiment 4 glossed sentences.

Learn the words. (下記の英単語(10語)を覚えなさい。)

| | | |
|---------------|--------|--|
| masco | 機関車 | Masco 50041 arrives in London at 4:15. He has designed a fast new masco . The driver got off the masco . |
| denent | 容貌 | Hers was the most beautiful denent he had ever seen. He was always careful to wash his denent . When he saw her denent it became whiter than ever. |
| hodet | 小道 | She walked slowly back down the hodet . The car was too large for the hodet so they walked in. He walked them out to the big car at the end of the hodet . |
| faddam | 住居 | My faddam is very far from here. He had decorated his new faddam with the finest furniture. Her faddam was the smallest on the street. |
| dangy | 巨石 | The dangy was as large as a small house. On the way up the mountain we passed a dangy . He stood on the biggest dangy to get a better view. |
| tasper | 熱望する | I used to tasper expensive clothes, and jewelry as well. Many women may tasper a baby. I do tasper a motor car. |
| cader | うとうとする | Some people cadered , but no one could sleep. Lying on the bed she had cadered . Closing her eyes she began to cader . |
| pacon | むせび泣く | She stopped screaming and began to pacon . After he left, she started to pacon . She fell onto the bed and started to pacon . |
| sagod | ひどく嫌い | His hobbies, which later he sagoded included dancing. I sagod the use of drugs in sports. We do sagod dust and dirt, and dirt in the bathtub. |
| ancon | 短剣 | She took a shiny ancon from her belt and pointed it at him. She picked up his ancon when he wasn't looking and cut him. She showed me how to hold the ancon . |

APPENDIX J

Experiment 4 productive writing task.

Write each word in a sentence in English and learn the words.

下記の単語を覚えなさい。それぞれの単語を使って英語で文章を作りなさい。

For example, man 男 I met a tall man yesterday.

| | | |
|--------|--------|-------|
| masco | 機関車 | _____ |
| denent | 容貌 | _____ |
| hodet | 小道 | _____ |
| faddam | 住居 | _____ |
| dangy | 巨石 | _____ |
| tasper | 熱望する | _____ |
| cader | うとうとする | _____ |
| pacon | むせび泣く | _____ |
| sagod | ひどく嫌い | _____ |
| ancon | 短剣 | _____ |

APPENDIX K

Experiment 5 treatment. Version 1.

Write each word in a sentence in English and learn the words.

下記の単語を覚えなさい。それぞれの単語を使って英語で文章を作りなさい。

For example, man 男 I met a tall man yesterday.

masco 機関車 _____

denent 容貌 _____

hodet 小道 _____

faddam 住居 _____

dangy 巨石 _____

tasper 熱望する _____

cader うとうとする _____

pacon むせび泣く _____

sagod ひどく嫌い _____

ancon 短剣 _____

Start the next page when you have finished writing all of the sentences.

すべての文を書き終えるまで次のページに進んではいけません。

(Experiment 5 treatment. Version 1 continued)

Read the sentences and learn the words. Do not write anything on this page.

下記の英文で下線をつけた単語の意味を考えなさい。このページに書き込んではいけません。

- mesut** なめる One of you can mesut the spoon and one can mesut the bowl.
The dog was trying to mesut the dish.
She mesuted her lips and moved her mouth.
- ictay** 槍 He was killed with the long hunting ictay.
He held a long ictay in his hand.
The easiest way to catch fish is with an ictay.
- tagon** 世捨て人 It was true that he was a tagon and never came to the village.
He had been a tagon, away from the world for three years.
He was not a tagon and travelled a lot in Europe.
- nuggy** 質に入れる He nuggied his watch to buy some new clothes.
The Queen nuggied her jewels to pay for his trips.
She went to the store and nuggied her wedding ring.
- nasin** ごみ廃棄場 Most of this garbage goes to nasin –massive holes in the ground far from where most people live.
The largest part of the nasins is paper and paper products.
950 old and dangerous nasins are to be cleaned up.
- copac** 嘆く They continue to copac for years after the death of their friend.
They still copac their dogs' death.
She thought of him, of how he must feel now, copacing his dead sister.
- hattaw** 修道院 She doesn't want to stay inside a hattaw for the rest of her life.
The Church of St. George was next to the hattaw.
The part of the hattaw in which I lived was called the noviceship.
- toncop** 埠頭 My brother and I were at the end of the toncop, fishing.
The Atlantic Ocean splashed against the toncop, and he stared out to sea.
He would have been happiest to stand on the end of a toncop and watch big ships go by.
- dapew** 礁 The small boat went south around the dapew.
I get very excited when I swim near a dapew.
With the dapew below, and a boat nearby to help them, they decided to try to go back the way they had come.
- gishom** マリネにする Pour brandy over the fruit and put it in a refrigerator for at least 12 hours to gishom.
He picked up his fork and looked at the small fish gishomed in olive oil and garlic.
Make the sauce in advance and gishom the meat for 2 hours.

Turn over the booklet when you have finished this page. このページが終わったら裏返しにしてください。

APPENDIX L

Experiment 6 receptive word pairs.

| | |
|--------|--------|
| mesut | 蛇口 |
| hodet | 小道 |
| denent | 槍 |
| dangy | 巨石 |
| nasin | 星座 |
| faddam | なめる |
| tasper | ひざをつく |
| copac | 水をかける |
| cader | うとうとする |
| pacon | むせび泣く |
| sagod | 礁 |
| dapew | 全速力で走る |
| ancon | 短剣 |
| gishom | 子馬 |
| masco | 機関車 |

APPENDIX M

Experiment 6 productive word pairs.

| | |
|--------|--------|
| 蛇口 | mesut |
| 小道 | hodet |
| 槍 | denent |
| 巨石 | dangy |
| 星座 | nasin |
| なめる | faddam |
| ひざをつく | tasper |
| 水をかける | copac |
| うとうとする | cader |
| むせび泣く | pacon |
| 礁 | sagod |
| 全速力で走る | dapew |
| 短剣 | ancon |
| 子馬 | gishom |
| 機関車 | masco |

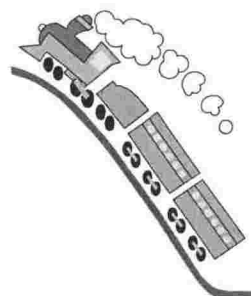
APPENDIX N

Experiment 6 picture description test.

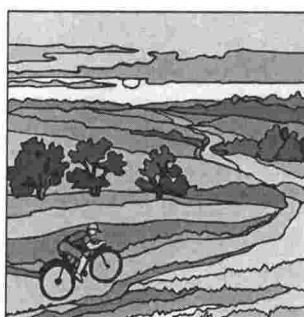
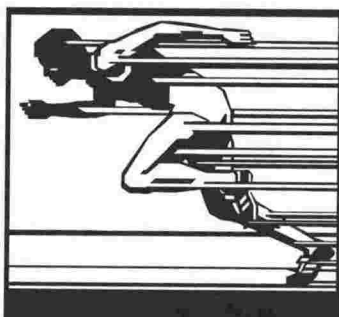
Write one sentence about each picture using a word that you learned today. 今日覚えた単語を使って、それぞれの絵についてひとつずつ文を作りなさい。



1. _____ 2. _____ 3. _____

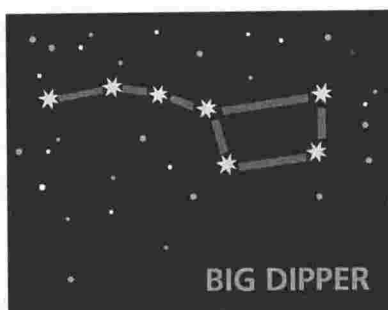


4. _____ 5. _____ 6. _____



7. _____ 8. _____ 9. _____

(Experiment 6 picture description test continued)

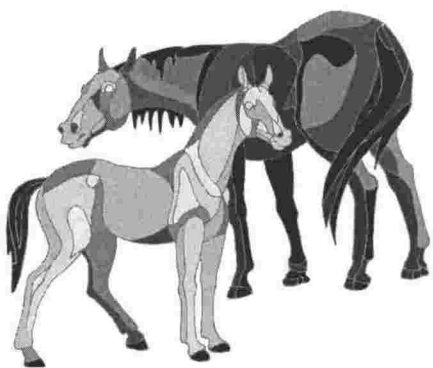


10. _____

11. _____



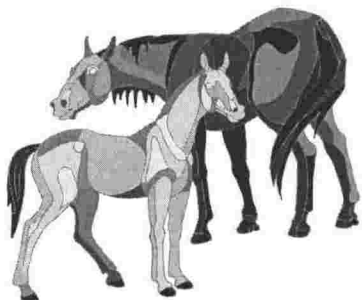
12. _____



13. _____



14. _____



15. _____



APPENDIX O

Experiment 6 true/false comprehension test.

Circle true or false for each sentence. それぞれの文について正しいか誤りに○をつけなさい。

- | | |
|---|------------|
| 1. People sometimes faddam their fingers. | True False |
| 2. People usually faddam a banana. | True False |
| 3. People often pacon when they are sad. | True False |
| 4. Men never pacon. | True False |
| 5. People sometimes cader on trains and buses. | True False |
| 6. Teachers are happy when students cader in class. | True False |
| 7. Parents sometimes tasper when they play with their children. | True False |
| 8. People are shorter when they tasper. | True False |
| 9. Ancons are heavier than computers. | True False |
| 10. Ancons are made of wood. | True False |
| 11. A dangy can be dangerous. | True False |
| 12. A dangy is soft. | True False |
| 13. Mascos are higher than houses. | True False |
| 14. Mascos are expensive. | True False |
| 15. Some sagods are famous. | True False |
| 16. Sagods are beautiful. | True False |
| 17. Hodets are bigger than roads. | True False |
| 18. Hodets are always straight. | True False |
| 19. Students sometimes brings denents to school. | True False |
| 20. Denents are very heavy. | True False |
| 21. You can touch a nasin but you cannot see one. | True False |
| 22. Nasins have names. | True False |
| 23. Gishoms are smaller than dogs. | True False |
| 24. A horse is a gishom when it is young. | True False |
| 25. You can copac water. | True False |

- 26. People sometimes copac sauce onto food.
- 27. People move slower when they dapew.
- 28. People sometimes dapew when they play sports.
- 29. Most houses in Japan have mesuts.
- 30. Mesuts are usually made of glass.

True False

True False

True False

True False

True False